# DOCUMENT RESUME

ED 267 078 TM 860 143

| | |
|---|---|
| AUTHOR | Wang, Margaret C., Ed. |
| TITLE | Temperament and School Learning. Papers presented at a Symposium at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984). |
| INSTITUTION | Pittsburgh Univ., Pa. Learning Research and Development Center. |
| SPONS AGENCY | National Inst. of Education (DHEW), Washington, D.C. |
| REPORT NO | LRDC-1985/25 |
| PUB DATE | 85 |
| NOTE | 146p. |
| PUB TYPE | Reports - Research/Technical (143) -- Collected Works - Conference Proceedings (021) |
| EDRS PRICE | MF01/PC06 Plus Postage. |
| DESCRIPTORS | *Academic Achievement; *Aptitude Treatment Interaction; Childhood Attitudes; Elementary Secondary Education; Emotional Disturbances; Higher Education; *Individual Characteristics; Individual Differences; Longitudinal Studies; Mexican Americans; Middle Class; Performance Factors; *Personality; Personality Measures; Psychological Characteristics; Self Control; Socioeconomic Influences; Student Adjustment; Student Behavior; *Student Characteristics; Teacher Student Relationship; Young Children |

ABSTRACT

Four research papers on temperament and school learning and two commentaries are presented. "Early Temperament and Later Educational Outcomes," by Jacqueline V. Lerner, Stella Chess, and Kathleen Lenerz, discusses temperament characteristics and academic attainment. One hundred thirty-three middle class subjects were studied from early infancy through adulthood, as part of the New York Longitudinal Study. Luis M. Laosa's paper, "Temperament, Performance, and Culture: Dimensions of Early Behavioral Style in Chicano Families," reports on a study of 100 Chicano children. Mothers rated temperament when their children were 2 1/2, 3, 3 1/2, and 4 years old. Mood/manageability and rhymicity/responsivity were the resulting components. Michael Pullis's research, described in "Temperament and Behavior Disorders: Teachers' Perceptions of Student Control," asked elementary and secondary teachers to rate 224 emotionally disturbed students' temperament and self-control. Results suggested that teachers' perceptions of students' self-control and socioeconomic status influenced their teaching styles. "An Analysis of Individual Differences in Student Temperament Characteristics and the Implications for Classroom Processes and Outcomes," by Lizanne DeStefano, Margaret C. Wang, and Edmund W. Gordon describes the instructional implications of student personality characteristics. Subjects ranged from age 4 to 9. Comments on the implications of all four papers are contained in discussant papers presented by Samuel Messick and Edmund W. Gordon. (GDC)

ED267078

# LEARNING RESEARCH AND DEVELOPMENT CENTER

1985/25

## TEMPERAMENT AND SCHOOL LEARNING

MARGARET C. WANG, EDITOR

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

J. Aug

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

University of Pittsburgh

TM 860143

# TEMPERAMENT AND SCHOOL LEARNING

Margaret C. Wang, Editor

Learning Research and Development Center
University of Pittsburgh

1985

The papers included herein were presented at a symposium held at the 1984 meeting of the American Educational Research Association, April, New Orleans. Preparation of the manuscript was supported by the Learning Research and Development Center, funded in part by the National Institute of Education (NIE), U. S. Department of Education. The opinions expressed do not necessarily reflect the position of NIE, and no official endorsement should be inferred.

# Temperament and School Learning

## Table of Contents

Preface
Margaret C. Wang ..................... v

Early Temperament and Later Educational Outcomes
Jacqueline V. Lerner, Stella Chess, and Kathleen Lenerz ..................... 1

Temperament, Performance, and Culture: Dimensions of Early Behavioral Style in Chicano Families
Luis M. Laosa ..................... 19

Temperament and Behavior Disorders: Teachers' Perceptions of Student Control
Michael Pullis ..................... 57

An Analysis of Individual Differences in Student Temperament Characteristics and the Implications for Classroom Processes and Outcomes
Lizanne DeStefano, Margaret C. Wang, and Edmund W. Gordon ..................... 81

Temperament Characteristics and Learning: Comments and Caveats
Samuel Messick ..................... 121

Temperament and Learning: A Commentary
Edmund W. Gordon ..................... 131

# PREFACE

The beliefs that all learners are alike in some ways yet unique along a number of dimensions, and that these unique individual characteristics differentially affect students' learning and adjustment, have come to be universally accepted. Furthermore, the educational implications of this common belief have been a continuing concern of researchers and practitioners alike. A basic premise underlying the work in this area is that effective schooling involves the design and implementation of learning environments which enable each student to achieve desired outcomes. Insofar as learning is a function of the learner's response to the environment, instruction is the intentional manipulation of the learning environment to facilitate appropriate responses. A major complicating factor in the purposeful design and use of the learning environment, however, is the diversity in learning characteristics of individual students and their requirements for achieving given outcomes.

Despite findings from psychological research that suggest a correlation between a variety of individual difference variables and learning, research on the educational implications of individual differences in learner characteristics traditionally has focused, for the most part, on status variables such as sex, ethnicity, social class, IQ, and level of academic achievement. Although these variables have proven to be useful in determining access, opportunity, and resource allocation, as well as in other aspects of educational policy and planning, they have had little influence on the design and implementation of educational practices. Among major reasons for this gap have been the lack of a descriptive base on a broad range of instructionally relevant individual difference variables and the lack of sufficiently detailed information that can be used in instructional planning and the actual instructional process. Temperament characteristics is an example of an individual difference variable for which a considerable empirical base has been amassed, yet very little work has been done to link the research base to educational practice.

The purpose of this monograph is to report the work of the few researchers who have pioneered an interest in examining the instructional implications of temperament characteristics of individual learners in school contexts. The papers focus specifically on the relationship between students' temperament characteristics and their school learning. Originally presented as a symposium at the 1984 Annual Meeting of the American Educational Research Association, the papers provide summaries and discussions of findings from several recent studies. The authors represent diverse perspectives, and the data reported in the papers were derived from a variety of contexts.

The first paper, entitled "Early Temperament and Later Educational Outcomes," is by Jacqueline V. Lerner and Kathleen Lenerz from The Pennsylvania State University and Stella Chess from New York University Medical Center. Based on data from parent interviews, direct observations, IQ testing, school interviews and observations, and clinical evaluations and treatment, the authors report findings on the role of temperament characteristics in predicting later educational outcomes. Temperament characteristics and learning outcomes from early infancy through adulthood were analyzed for 133 subjects. Particular attention was paid to the implications for the use of information on differences in temperament characteristics to shape the nature of relationships and classroom interactions between teachers and individual students.

The second paper in this monograph is by Luis M. Laosa from the Educational Testing Service in Princeton, New Jersey. Entitled "Temperament, Performance, and Culture: Dimensions of Early Behavioral Style in Chicano Families," the paper reports findings from a short-term longitudinal study designed to investigate the population validity of current conceptions of temperament. Data for a sample of 100 Chicano families at widely varied socioeconomic levels were analyzed to identify the dimensions of temperament perceptions that accounted for variance in mothers′ responses to the Behavioral Style Questionnaire, and to determine whether variations along these dimensions could be predicted by measures of children′s and mothers′ cognitive competence and family background characteristics. Variations along specific dimensions of mothers′ temperament perceptions were found to be related to the degree of mother tongue maintained in the study′s bilingual and bicultural sample of families, as well as to children′s performance on tests of cognitive ability and achievement.

The third paper, entitled "Temperament and Behavior Disorders: Teachers′ Perceptions of Student Control," is by Michael Pullis from the University of Texas at Dallas. The study reported in this paper focused on examining the relationship between teachers′ perceptions of the amount of control evidenced by the sample of emotionally disturbed students over their classroom behavior and teacher-student interactions as reflected in measures of classroom decision making, aspects of disordered behavior, and teacher stress. The findings lend some support for the notion that teachers′ perceptions concerning the amount of control that individual students have over their particular behavior problems may be an important mediating factor in the nature of teacher-student relationships, specific instructional strategies, and teachers' classroom decisions.

The final paper in this monograph is by Lizanne DeStefano and Margaret C. Wang from the University of Pittsburgh and Edmund W. Gordon from Yale University. Entitled "An Analysis of Individual Differences in Student Temperament Characteristics and Their Implications for Classroom Processes and Outcomes," the paper reports the findings from a descriptive study designed to investigate the interactive effects of temperament characteristics, school learning environments, student behavior, and student learning outcomes. Results from analysis using triangulation procedures, observational data, and student and teacher ratings suggest several patterns of distinct relationships among these clusters of variables. The authors draw implications from their findings for adaptation of individual student temperament characteristics to the requirements of particular learning environments, and for the incorporation of information on differences in student temperament into schools' program design and implementation efforts.

Discussant papers by Edmund Gordon from Yale University and by Samuel Messick from the Educational Testing Service address the implications of findings reported in the four papers as they relate to the state of the art of research on temperament and learning and to the provision of educational environments that are responsive to differences in student temperament characteristics.

The underlying assumption of all the papers in this monograph is that individual differences in temperament characteristics are related to differences in student learning outcomes. Intervening variables in this relationship are identified as including teacher and family perceptions of student temperament, cultural background, teacher-student interactions, instructional strategies, and the design of the learning environment. By addressing and describing the dynamic and multifaceted interaction of individual differences in temperament characteristics with the classroom learning process, the authors have provided a springboard for further research and program development aimed at fortifying the link between the state of the art and the state of practice.

In addition to acknowledging the contributions of all the authors, I would like to express my appreciation to Ed Gordon for his encouragement and support in initiating this project and to Rita Catalano and Regina Rattigan for their editorial assistance in preparing this monograph.

Margaret C. Wang
Pittsburgh, Pennsylvania
Summer, 1985

# EARLY TEMPERAMENT AND LATER EDUCATIONAL OUTCOMES

Jacqueline V. Lerner
The Pennsylvania State University

Stella Chess
New York University Medical Center

and

Kathleen Lenerz
The Pennsylvania State University

The focus of much of the research on children's educational adjustment and achievement has been on cognitive and intellectual factors and how these factors vary and are related to differential educational outcomes. Although these factors are important, researchers have recently raised concerns that other sources of individual differences and how they relate to educational outcomes are in need of more intensive study (Keogh & Pullis, 1980). For example, some have argued for the inclusion of affective measures when predicting academic performance (Kohn, 1977; McKinney, Mason, Perderson, & Clifford, 1975; Perry, Guidubaldi, & Kehle 1979); others have suggested the need to study the behavioral stylistic aspects of achievement related behaviors (Thomas & Chess 1977). These stylistic aspects of behavior (or temperament) refer to the how of behavior; not what a child does, but how he/she goes about doing it. Thomas and Chess have identified nine dimensions of temperament in their 27-year-old New York Longitudinal Study (NYLS), which began as a study of individual differences in children. These nine dimensions are activity level, rhythmicity, approach/withdrawal, adaptability, intensity of response, threshold of responsiveness, quality of mood, distractibility, and attention span/persistence. Thomas and Chess (1977) argue that these behavioral style characteristics can have a strong influence on both the parent child and teacher-student relationships. For example, problems in development can occur when there is a conflict between the child's behavioral style and the parents' demands or reactions to it. In the same way, teacher-student conflict can arise and in turn, have strong effects on school performance.

Teachers' evaluations of and reactions to student characteristics can serve to either enhance or negatively affect school performance. Gordon and Thomas (1967) have emphasized this in a study of the relationship between teachers' perceptions of student temperament and their estimates of student intelligence. Teachers were more likely to underestimate the intelligence of children who were slow in approaching and adapting to new situations. Gordon and Thomas argue that these underestimates of ability may affect school performance because teachers' expectations may be lower than the students' actual abilities, thus negatively affecting the type and level of educational activities and opportunities offered the students.

Other research on the relationship between children's temperament and educational factors supports the notion that temperament may indeed be a salient factor in a child's educational achievement and adjustment For example, there is evidence for a relationship between temperament and ratings of educational competence and educational risk (Hall & Keogh, 1977). In addition, Pullis and Cadwell (1982) found that teachers used information about their students' temperament characteristics more than any other type of information when making classroom management decisions.

In sum, the above findings converge in suggesting that temperamental characteristics play an important role in teacher-student relationships and may relate to educational outcomes. The present study investigated the role of temperamental characteristics in the prediction of later educational outcomes in the children of the NYLS (Thomas & Chess, 1977; Thomas, Chess, & Birch, 1968; Thomas, Chess, Birch, Hertzig, & Korn, 1963). Specifically, temperamental attributes of the children at ages 3 and 5, as well as the ratings of the Easy and Difficult Child clusters (Thomas & Chess, 1977) were used in the prediction of educational competence both at adolescence and early adulthood.

On the basis of findings from studies independent of the NYLS that suggest that the presence of Difficult temperament is associated with various problems related to both home and school (Carey, 1972, 1974; Graham, Rutter, & George, 1973), and from findings within the NYLS sample that link the Difficult temperament to behavior problems (Thomas & Chess, 1977; Thomas, Chess, & Birch, 1968), it was expected that Difficult temperament at ages 3 and 5 would relate to lower educational achievement outcomes in adolescence and adulthood.

## Method

### Subjects

The sample consisted of the 133 white middle-class children (66 males, 67 females) from the New York Longitudinal Study (NYLS) founded by Thomas and Chess in 1956. The subjects were 78% Jewish, 7% Catholic, and 15% Protestant. They were all from New York City and surrounding areas. Data collection began with an initial parental interview (during the first months of each child's life) and continues to the present day. The sample has remained intact throughout the study and a wealth of data exists for each subject. The methods of data collection used in this almost 30 year longitudinal study have been described repeatedly in the literature (Thomas et al., 1963; Thomas & Chess, 1977; Thomas, Chess, & Birch, 1968). Only the measures pertinent to the focus of the present study will be summarized here.

## Measures

Three-year and five-year temperament ratings. Parent interviews at ages 3 and 5 contained detailed information on the characteristics of the children's behavior. It was possible from these interviews to rate each child's temperament or behavioral style along each of the nine NYLS dimensions.

The nine temperament dimensions were defined and rated as follows: activity level -- the motor component present in a child's functioning, scored as low, moderate, or high; rhythmicity -- the regularity of a child's functions such as sleep-wake cycles, hunger, feeding patterns or elimination, scored as regular variable, or irregular; approach/withdrawal - the nature of the initial response to a new stimulus, scored as approach, variable, or withdrawal; adaptability -- ease with which a child adjusts to new or altered situations, scored as low, moderate, or high; threshold of responsiveness -- the intensity level of stimulation needed to evoke a response, scored as low, moderate, or high; intensity of reaction -- the energy level of a response, irrespective of its quality or direction, scored as mild, variable, or intense; quality of mood -- the amount of pleasant or unpleasant behavior, scored as positive, variable, or negative; distractibility -- the effectiveness of extraneous stimuli in interfering with or altering the direction of the ongoing behavior, scored as nondistractible. variable, or distractible; and attention span/persistence -- attention span refers to the length of time a child pursues a particular activity; persistence refers to the continuance of an activity in the face of obstacles, scored as non-persistent, variable, or persistent.

Easy/Difficult temperament scores. At each age it was also possible to obtain an "Easy" or "Difficult" temperament score for each subject. Specifically, each subject received a score which indicated the amount of Easy (positive mood, high rhythmicity, low or mild intensity of reactions, high adaptability, positive approach) versus Difficult (low rhythmicity, negative withdrawal responses, slow adaptability, negative mood, high intensity of reactions) temperament characteristics they possessed. The present research included only the nine individual temperament scores and the Easy/Difficult scores for ages 3 and 5 in the analyses. The focus on temperament at ages 3 and 5 is prompted by the fact that these years have been found to be particularly salient in the prediction of later outcomes in previous research with the NYLS sample (Lerner & Vicary, 1984; Thomas & Chess, 1977).

Adolescent achievement ratings. The Wide Range Achievement Test (WRAT) was given to 56 subjects when they were between 10 and 15 years of age. Individual WRAT reading, spelling, and arithmetic scores were used as the outcome measures of achievement in adolescence.

Early adult achievement. Objective tests and/or college grades were not consistently available for all subjects in early adulthood. Therefore, as a measure of educational outcomes in early adulthood, each subject was rated on his/her reported level of educational attainment as of the young adult interview (subjects were approximately 19-22 years old). Ratings (from 1-5) were as follows: high school graduate, some college or technical school, B.S. or equivalent, M.S. or equivalent, and Doctorate or equivalent (includes M.D. or law degree), respectively.

## Analyses

To address the question of whether early temperamental characteristics play a role in the prediction of adolescent and early adult educational outcomes, a series of multiple regression analyses were used. Because the sample size was only 56 for the analyses using the WRAT scores, only five predictors could be used. Thus, the five individual temperamental characteristics which make up the Difficult Child cluster were used as the predictors of WRAT reading, spelling, and arithmetic scores. This was done first using the age 3 temperament ratings and then using age 5 temperament ratings (which yielded a total of six analyses). Next, the remaining four temperamental characteristics of activity level, threshold of responsiveness, distractibility, and attention span/persistence were used to predict the WRAT scores of reading, spelling, and arithmetic. This was done separately for ages 3 and 5, thus yielding another set of six regression analyses. In addition, another series of multiple regression analyses were done using ages 3 and 5 Easy/Difficult scores as the predictors of WRAT reading, spelling, and arithmetic scores, which yielded a total of three analyses.

These same predictors (the five Difficult child attributes; the four non-Difficult attributes) at ages 3 and 5 and the Easy/Difficult scores at ages 3 and 5 were then used to predict the criterion of educational attainment in early adulthood, which yielded a total of six analyses.

## Results

No significant relationships between early temperament and early adult educational attainment were obtained for the present sample. This was the case both when individual temperament scores and Easy/Difficult composite scores were used as predictors.

However, in adolescence a different picture emerged. When individual temperament attributes were used as predictors, 33% of the analyses were significant; however, no systematic relationships were obtained. However, when the composite Easy/Difficult scores at ages 3 and 5 were used as predictors, significant findings were obtained. Specifically, the Easy/Difficult child composite score at age 5 significantly predicted all three WRAT scores (Reading: $R^2 = .17$, $p < .002$; Spelling: $R^2 = .10$, $p < .02$; Arithmetic: $R^2 = .11$, $p < .01$). In all three cases, higher difficulty scores were related to lower levels of achievement. These relationships were maintained even after IQ was controlled. In one case, for the WRAT reading scores, the Easy/Difficult score at age 3 was a significant predictor, and indicated that higher difficulty was related to higher reading scores. However, because age 3 Easy/Difficult scores were not related to the other achievement scores, and because the relationship was low ($R^2 = .07$, $p < .04$), we view this as more of a possible random occurrence than of a true indication that easier children at age 3 have lower reading scores in adolescence.

## Discussion

The results of the present study lend some support to the notion that temperament may play a role in the development of educational competence. Nevertheless, we did not find that early temperament bore any relationship to our measure of early adult educational attainment. This may have occurred for a number of reasons. First, there was limited variability in this particular measure, since most of the sample were in college at the time of measurement, and thus may have restricted the power to detect a relationship if one existed. Second, Difficult temperamental characteristics may disrupt teacher-student relationships in the early years, hence, influencing academic *performance*. However, our adult measure was one of *attainment*; performance was not directly measured. In addition, this sample was a middle-class one, with an average IQ of 123, hence, the opportunities for high educational attainment were probably consistent and equally expected across the entire sample. We do not mean to say here that Difficult temperament should *not* bear any relationship to adult educational outcomes. Certainly if a person remained Difficult throughout their life it could disrupt their functioning in many ways.

However, the stability of the Difficult temperament from childhood to adulthood is not high (r = .15 from age 5 to adulthood, and r = .31 from age 3 to adulthood), which may indicate that as the person matures and his/her interactions become more diffused across various contexts, the presence of a Difficult temperament may not have such a strong impact. In childhood, interactions may be more intensely focused in one or a few contexts, since the child spends most of his/her time with parents or teachers. Thus, the presence of a Difficult temperament may interfere more with parent-child and student-teacher interactions.

The present findings do indicate, on the other hand, that the presence of a Difficult temperament at age 5 is predictive of lower levels of achievement in adolescence, although difficulty at age 3 was not predictive of later academic competence. We can speculate as to why age 5 difficulty is a more salient predictor than age 3 difficulty. The child at age 5 is presented with many more new sets of behavioral demands than he/she may have encountered at age 3. Parental demands for more mature behaviors may increase at this time, along with demands that arise from the school situation. Thus, in order to be seen as Difficult at age 5, the child is probably manifesting Difficult behaviors in more spheres of functioning than the 3-year-old child, leading to a stronger indicator of Difficulty.

The implications of the relationship between Difficult temperament and achievement are twofold. First, although we cannot assume that the causality of this relationship lies in the influence it has on teacher-student interaction, it is not unlikely that the presence of the Difficult Child characteristics may indeed influence teacher-student interaction. As we have noted above, Gordon and Thomas (1967) found that teachers were influenced by children's temperamental characteristics, and were more likely to underestimate the ability of children who showed low approach behaviors and slow adaptability. In addition, Pullis and Cadwell (1982) found that teachers used information about their students' temperamental characteristics in making classroom management decisions. These findings have concerned only one or two of the temperamental characteristics. We may speculate here that the combination of all five Difficult Child characteristics (high irregularity, withdrawal responses, slow adaptability, negative mood, high intensity of reactions) may indeed impact on the relationship between the child and the teacher, creating negative feedback to the child, fostering a lowered self-concept and sense of defeat, the underestimation of the child's true abilities, and hence, decreased performance.

The second implication of our results is that, if our findings and those of others can be taken and, hence, *can* influence the student-teacher relationship and hence affect performance, then steps should be taken to inform teachers of these possible influences, to make them aware of individual differences in their students' characteristics and how they might become more accommodating to them. Lewis (1977) suggests that information about children's temperamental characteristics could serve as a useful tool for teachers so they could anticipate a child's reaction and make modifications to facilitate rather than impede the child's adjustment and progress. We too feel that these results and others like them should be taken as an impetus to sensitize parents and teachers to the nature of individual differences in children's temperament and how they can influence the child's functioning.

## Acknowledgment

Jacqueline V. Lerner's work on this paper was supported by a grant from the John D. and Catherine T. MacArthur Foundation. The authors thank Richard M. Lerner for his critical reading of an earlier version of this paper.

20

## References

Carey, W. B. (1972). Clinical applications of infant temperament measurements. *Journal of Pediatrics*, *81*, 823-828.

Carey, W. B. (1974). Night-waking and temperament in infancy. *Journal of Pediatrics*, *84*, 756-758.

Gordon, E., & Thomas, E. (1967). Children's behavioral style and teachers' appraisal of their intelligence. *Journal of School Psychology*, *5*(4), 292-300.

Graham, P., Rutter, M., & George, S. (1973). Temperamental characteristics as predictors of behavior disorders of children. *American Journal of Orthopsychiatry*, *43*, 328-339.

Hall, R., & Keogh, B. (1977). Qualitative characteristics of educationally high-risk children. *Learning Disabilities Quarterly*, *1*(2), 62-69.

Keogh, B., & Pullis, M. (1980). Temperament influences on the development of exceptional children. In B. K. Keogh (Ed.), *Advances in special education* (pp. 239-279). Greenwich, CT: JAI Press.

Kohn, M. (1977). *Social competence, symptoms, and underachievement in childhood: A longitudinal perspective*. Washington, DC: V. H. Winston & Sons.

Lerner, J. V., & Vicary, J. R. (1984). Difficult temperament and drug use: Analyses from the New York Longitudinal Study. *Journal of Drug Education*, *14*, 1-8.

Lewis, J. L. (1977). The relation of individual temperament to initial social behavior. In R. D. Smart & M. S. Smart (Eds.), *Readings in child development and relationships* (2nd ed.) (pp.76-94). New York: Macmillan.

McKinney, J., Mason, J., Perderson, K., & Clifford, M. (1975). Relationship between classroom behavior and academic achievement. *Journal of Educational Psychology*, *67*, 198-203.

Perry, J., Guidubaldi, J., & Kehle, T. (1979). Kindergarten competencies as predictors of third-grade classroom behavior and achievement. *Journal of Educational Psychology*, *71*, 443-450.

Pullis, M., & Cadwell, J. (1982). The influence of children's temperament characteristics on teachers' decision strategies. *American Educational Research Journal*, *19*, 165-181.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Thomas, A., Chess, S., & Birch, H. (1968). *Temperament and behavior disorders*. New York: New York University Press.

Thomas, A., Chess, S., Birch, H., Hertzig, M., & Korn, S. (1963). *Behavioral individuality in early childhood*. New York: New York University Press.

## Case Vignettes

The following vignettes were chosen from among the NYLS sample to illustrate some points with respect to the findings of the present study. The descriptions of the children presented are not what would be expected based on our findings that Difficult children were more likely to suffer from academic problems than Easy children. However, we wanted to illustrate that even though it is more probable statistically that Difficult children will have more achievement difficulties than Easy children, individual outcomes are a result of the interaction of temperament in specific contexts. Therefore, the outcomes for both Difficult and Easy children could be quite different than expected, as illustrated below. We believe that these case studies support the idea that if parents and teachers are sensitive to the individual characteristics of the child, difficulties can be overcome and positive outcomes can result.

### Dorothy

Temperament characteristics were slow moving motorically, mood negative but near the neutral line; rhythmicity variable; intensity low; withdrew to new situations; adapted slowly; threshold moderate; very low distractibility; persistence and attention span depended upon the situation with very high persistence and very long attention span for self-selected problems which were often quite difficult, but, if not motivated, both persistence and attention span were markedly low.

Example at age 4: wished to learn to sew on a button -- was told it was too hard. She insisted, watched a token exhibition, then practiced for days and mastered it.

In school, each new learning task presented difficulties and whining complaints, then was mastered with persistent effort, then enjoyed with quiet self-given tasks but little other expression of satisfaction. After second grade, there was a school change. Because of marginal performance and the fact that her age was suitable for either grade, the new school suggested that she repeat second grade and achieve a strong educational base. Dorothy would not, saying that she had "done" second grade already, adding that if she didn´t try third, how could she know whether she could "do" it, and promising not to fuss if it was then decided to put her back. She had a difficult beginning, persisted, and mastered the grade in the same style as always.

Dorothy had several sisters, one of whom was notably quick moving, who caught on rapidly to new learning. Throughout all school years, Dorothy complained about her slow pace; that she had to work so hard to learn while her sisters achieved equally good grades with much less effort. Nevertheless, her level of mastery was high as attested to by the top level college into which she was accepted, and from which she graduated with respectable grades -- still in her plodding manner.

In her chosen field, as a young adult, Dorothy works in the same manner, mastering each new technique with slow persistence and feelings of ineptitude, but acknowledging with each mastery that her level of competence was high -- as did those who were capable of judging her work.

Carl

Temperamental qualities were those of the Difficult child, on the extreme end of the continuum. Rhythmicity was irregular, mood negative, intensity of expression high, had high withdrawal from new situations; adapted very slowly.

Other temperamental qualities included moderate threshold level, average activity level, moderate distractibility, high persistence and long attention span.

After a few weeks of nursery school, parents were advised by the school that Carl was "not ready" for nursery, and the director suggested that he wait until the following year. Since this type of thing had been occurring from infancy, namely that Carl's initial reaction had been a tumultuous rejection of each new food, place, activity, person -- with very slow familiarization and almost universal final acceptance and zestful involvement -- after consultation with us, the parents declined to withdraw Carl from nursery. By Thanksgiving time, he was looking forward to school each day and being most helpful in the school routines such as handing the juice around and getting out the cots for rest time.

This scenario was the template for Carl's early school years. Each fall, with a new teacher, some new students, and the change from the vacation organization of life, Carl had an initial period of stormy withdrawal, slow adaptation, and finally academic and social mastery accompanied by high intensity of mood expressiveness. Once this mood became positive in orientation, the term changes from "tantrum" to "zestful." Since the family did not move throughout Carl's childhood, changes in school were minimal -- the building remained the same, schoolmates and friends were largely constant; while teachers changed from year to year, they were not unknown by reputation. With progress to junior high school and senior high school, the same group of children moved together. Thus, the environment was relatively stable, learning tasks altered gradually, and social companions were almost constant.

Other than a brief period of discomfort each school year start, basically Carl was known as a high achiever, a welcome companion, be it sports or class work, committee, or cultural group interest.

With the first year of college, Carl had a surprising (to him) experience. He had selected a college to which none of his friends were going, thinking nothing of this circumstance. Once there, he found that he was uneasy with other students, found the change in work approaches and need to plan differently most difficult to assimilate -- for example, planning ahead for several large reports all due at the end of the term in different subjects. Most of all, he had the feeling "this isn´t me." His unease and unhappiness were expressed with vigor to his family, who suggested that, when home next, he have a discussion with one of the psychiatrists on the NYLS project. By the time of this interview, Carl had, on his own, taken steps which were well chosen. He had lightened his academic program, which had been in fact overloaded due to his own prior experience that he had managed an overly large academic program with ease in high school, and due to his desire to try out many subjects early so as to have a basis for selecting his major. Instead, he found that each subject felt unfamiliar and even painfully so. He had moved from his dormitory into an apartment where he would be forced by circumstances to live in proximity to others, and forced himself to attend social occasions even though he felt ill at ease. He timed himself while studying in the library, accepting nothing less than a two-hour stint. He took one cultural extracurricular task, the clavicord, for pleasure, reasoning that, as he was a good pianist, this new technique of a keyboard would give him pleasure. In fact, he had achieved a number of acquaintances, some of whom he could envisage as friends; he was now on top of his academics and mastering them with no need to time his studying, and he felt more like "myself."

He was asked about his initial introduction to the piano, in middle childhood. A friend had started piano lessons and Carl had asked his parents to allow him to do so also. They started the lessons after extracting from him the promise that he would continue with them for six months no matter how he felt. Very soon he was hating them but was held to his promise. When, after six months, his mother asked if he wished to stop, Carl stated his enjoyment and his desire to continue and indeed had forgotten some time before that he had ever wished to stop. He had forgotten this initial withdrawal reaction to piano and was in fact making good progress with the clavicord.

We asked ourselves whether, in the years between the early uncompromising expression of the Difficult Child and the end of high school, Carl's temperament had changed, or, like an underground stream, had submerged for a time. On reflection, it would appear that on the contrary, we had here an example of the interactive aspects of temperament -- of the indication that temperamental expression occurs in response to the environmental demands. Once familiarization had been achieved, it had always been characteristic of Carl that he demonstrated positive mood with intensity of expressiveness. During the years of middle childhood, the circumstances of his life had in fact brought few new demands. When these came on the scene, as with piano lessons, an initial stormy withdrawal was still characteristic. The circumstances at college had brought together many new demands and expectations. However, Carl's self-image was that of having many friends, of easy mastery of tasks that were difficult, of having zest as his companion. With this self-image, the motivation to become himself again led to his own appropriate solution. After discussing the temperamental issues, Carl was given his nursery school data to read, which he did with merriment. His comment was that, whenever he enters a new situation in the future, he will know what to do and will remember that the discomfort will be limited.

## Kate

Temperamentally, Kate was an easy child. She presents an illustration of the qualitative circumstances in which one of the subjects went against the quantitative flow. Temperamentally, Kate was regular in rhythm, showed positive mood, was moderately intense in mood expressiveness, approached new situations, adapted swiftly. She was also average in both threshold and activity level, and had average distractibility but low persistence and attention span.

As an infant and in early nursery school, comments by teachers were positive. But as rules and restrictions were introduced, more and more we were told in the teacher interviews that Kate seemed to turn all situations into social interchanges. This was often appropriate, but at other times when instructions were given or tasks assigned, Kate seemingly ignored the fact that she was expected to address herself to them but instead chattered on happily. While Kate had been very much liked by the other children, as kindergarten, then grades entered her life, she was falling behind academically despite above-average intelligence. Further, although socially inclined and ready to join in others' ideas, the other children began to reject her since Kate seemed not to follow game rules or agreed-upon plans -- despite the fact that she had no objection to them; rather, she seemed oblivious to anything except social interchange.

Discussion with parents was revealing. Each had been a good student, and each felt that family had been unduly pressuring. A good mark was only good enough if it was perfect; any accomplishment was met with pressure to undertake another harder and higher goal. With such perceptions of their own pressured childhoods, the parents had decided that Kate would not have similar experiences. They therefore reacted, when Kate brought home news of accomplishments, to the fun the child had had, and downgraded either by lack of attention or by so saying, that the fact of task accomplishment was not of value. There had been some instances when mother had been annoyed by Kate's sublime indifference to a home rule, but being a "yeller" according to her own self-characterization, mother simply yelled a bit, then failed to carry through. Thus, Kate never understood what the issue had been. With this constant reinforcement of the social aspects of interaction being all important, Kate did not understand what was at stake in school when teachers told her to get on with the task -- not having heeded instruction, she didn't know what she was supposed to get on with. Nor, when classmates became annoyed with her failure to follow rules or do her bit of a committee homework assignment, did Kate understand why they had, seemingly for no reason, started not to like her.

This was discussed with parents, who understood the issue intellectually. However, when follow-up was done a year later, although the discussion was recalled correctly as to the problem, timing of its occurrence, and advice given, both mother and father were in agreement that they had in fact not altered their message to Kate. Mother said hopefully, "Maybe she will catch on in school. So she'll be beautiful but dumb."

And indeed, Kate's academic and social progress has been marginal. Close to both parents, she adapted to their values.

# TEMPERAMENT, PERFORMANCE, AND CULTURE: DIMENSIONS OF EARLY BEHAVIORAL STYLE IN CHICANO FAMILIES

Luis M. Laosa

Educational Testing Service
Princeton, New Jersey

In recent years there has been a marked resurgence in the attention given to temperament as a construct in the analysis of human behavior and development. Despite the growing interest in identifying and assessing temperamental characteristics, little effort has been directed at studying temperament in ethnic minority populations. A recent review of approaches and techniques for assessing temperament indicates that standardization and normative samples represent, with only a few exceptions, a U. S. middle-class White population (Hubert, Wachs, Peters-Martin, & Gandour, 1982). This state of affairs poses serious limitations to our knowledge because of the risks involved in generalizing research findings beyond the sociocultural population sampled in any given study.

Of the various formulations of temperament, perhaps the most influential is the one derived from the New York Longitudinal Study (NYLS) (Thomas & Chess, 1977; Thomas, Chess, & Birch, 1968; Thomas, Chess, Birch, Hertzig, & Korn, 1963). There is no agreement as yet, however, on a generally accepted definition of temperament. Disagreements revolve around differing assumptions as to the nature, origin, and constancy of temperament. Moreover, diverse theoretical orientations have provided somewhat different formulations as to the number, specificity, and stability of the hypothesized categories or dimensions of temperament. Nevertheless, most formulations of temperament assume that significant and relatively enduring differences among individuals can be identified early in the stylistic aspects of behavior. Temperament is most often measured through parental report (questionnaire, interview); thus, parental perceptions are heavily implicated in temperament research. Temperament -- and perceptions of it -- have been hypothesized to be related to other characteristics of the child (e.g., cognitive competence) and to characteristics of the child's family and social environment, although almost always in ways that are either vaguely specified or difficult to operationalize (for reviews see Bates, 1980; Campos, Barrett, Lamb, Goldsmith, & Stenberg, 1983; Keogh & Pullis, 1980; Lerner & Lerner, 1983; Rutter, 1982).

Questions of construct validity in general and, more recently, issues regarding a special case of construct validity, namely population validity, surround current conceptions of temperament. The concept of population validity expands the concept of construct validity by implying that an inference may be valid for one sociocultural population and not for another. Population validation is the process of discovering the meaning that a measure or combination of measures has when applied to a population of interest.

To illustrate that the sociocultural context can influence the functional meaning of temperament, the NYLS researchers compared their core sample, which consists predominantly of professional Jewish parents, with a sample of low-income Puerto Rican families also residing in the New York City area. The following difference in socialization practices and their sequelae are especially relevant. The professional Jewish parents generally made more demands for their children's early establishment of regular cycles (rhythmicity) for such biological functions as sleep-wake and feeding schedules than did the low-income Puerto Rican parents. Consequently, arhythmicity of biological functions as a temperamental characteristic in infancy and early childhood was a major source of adjustment problems when it occurred in the core sample, but not so when it occurred in the low-income Puerto Rican sample. Later, however, when the children entered school and had to face its inflexible system of schedules, arhythmicity became a source of adjustment problems also in the low-income Puerto Rican sample (Thomas, Chess, & Korn, 1982; Thomas, Chess, Sillen, & Mendez, 1974). This observation touches on a central aspect of population validity, namely, that the specific behaviors used as markers for constructs may vary in their meaning and functional significance across sociocultural contexts (Laosa, 1977, 1979, 1981a, 1981b).

The analyses reported here examined mothers' perceptions of their young children's temperament as measured by the Behavioral Style Questionnaire (BSQ) in a sample of Chicano families of widely varied socioeconomic levels. The BSQ, designed and constructed on the basis of the NYLS conceptual framework (McDevitt & Carey, 1975, 1978), is probably the most widely used technique for assessing temperament in young children. Chicanos comprise one of the largest and fastest growing ethnolinguistic groups in the United States; there are, however, to the writer's knowledge, no previous published data on temperament in Chicano samples.[1]

Two questions were addressed in the analyses now reported -- "What dimensions account for variance in the Chicano mothers' BSQ responses?"; and

"Can variations along these dimensions be predicted from characteristics of the child, the mother, and the environment?" The objective was twofold -- first, to identify in the Chicano mothers' reponses to the BSQ items those dimensions that showed short-term continuity and stability (Emmerich, 1968); and second, to determine whether variations along these dimensions of temperament perceptions could be predicted in part by (a) measures of the children's abilities and achievement, (b) measures of the mothers' intelligence, and (c) family background characteristics that are considered important aspects of demographic diversity within this ethnolinguistic group.

## Method

### Participants

The focal participants were 100 Chicano children (44 boys, 56 girls) and their respective household families. No two focal children belonged to the same household. The sample was followed longitudinally; a battery of measures was administered at ages 2 1/2, 3, 3 1/2, and 4 years -- the Behavioral Style Questionnaire was included at the latter two points. The sample included only households in which both of the child's parents identified themselves and both of their respective parents by such terms as "Mexican American," "Mexicano," "American of Mexican origin," or "Chicano."[2]

To obtain the pool from which the participants were selected, invitations for potential participation in the study were mailed in both Spanish and English to all Spanish-surnamed parents registered as having given birth to a child in Bexar County (San Antonio), Texas from October 15, 1975 through July, 1976. In addition, invitations were published in newspapers and posted throughout the city in supermarkets, pediatricians' offices, churches, schools, and other places. On the basis of information obtained from both (a) the forms in which parents gave consent to be considered as potential participants, and (b) subsequent interviews conducted either by phone or through a home visit, the families who consented to be considered were screened, and the sample was selected to be as representative as possible of the Chicano population in the United States with regard to the distributions of parents' schooling level and household income. To control for the possible confounding effects of variations in parental absence/presence, maternal employment/unemployment, and out-of-home care, the sample included only households in which (a) both parents lived at home with the child, (b) the mother was not employed outside the home,[3] and (c) the child did not attend a nursery school or day care center. In addition, to minimize the possibility of longitudinal attrition, preference was given to the selection of (a) families with a long history of residence in south-central Texas (families with plans to move from this area during the period of data collection were excluded from the sample); and (b) families in which the child's father was steadily employed. Accordingly, 91% of the mothers and 85% of the fathers were born in the continental United States; and all the fathers except three were employed. At the last longitudinal point of data collection, 95% of the families were still intact, 29% of the mothers were employed outside the home (part-time, 9%; full-time, 20%), and 8% of the children were attending a nursery school or day care center. Children with known physical or mental abnormalities were excluded from the sample. Each of the selected families received a modest honorarium for their participation in the study.

The mean age of the parents was 28.7 for mothers (SD = 5.0) and 30.9 for fathers (SD = 6.0). Most of the families (90%) were Catholic; the remainder reported various Protestant affiliations. Table 1 presents additional data descriptive of the sample.

It is helpful to compare the characteristics of the present sample with statistics reported by the U. S. Bureau of the Census. The median household (cash) income of the present sample was $10,025 -- very similar to that of all husband-wife families (wife not in paid labor force) in the total U. S. Mexican-origin population, $10,914 (U. S. Bureau of the Census, 1979). The median school years completed by the fathers and mothers in the present study was 12.0 and 11.6 -- almost identical to the total U. S. Mexican-origin population in the age range of 25 to 34 years, 12.0 for men and 11.5 for women (U. S. Bureau of the Census, 1979). In sum, the sample is very similar to the total U. S. Chicano population in its distribution on both parental schooling level and income.

## Temperament Measures

The Behavioral Style Questionnaire (BSQ) was administered to each mother twice: initially when the child was 3 1/2 years old (plus or minus 2 months) and again 6 months later (plus or minus 2 weeks). The BSQ is designed to yield measures of the nine NYLS categories of temperament in 3- to 7-year-olds (McDevitt & Carey, 1975, 1978). It consists of 100 items, each describing a specific event. (Item 21 was deleted because it applies only to children

Table 1

Labels, Descriptions, and Descriptive Statistics of Family Background, Home Language, and Performance Variables

| Variables | N | M | SD |
|---|---|---|---|
| Family Background | | | |
| Household income (annual dollars) | | | |
| At age 2 1/2 [a] | 100 | 11,171.1 | 5,633.7 |
| At age 4 | 93 | 14,245.1 | 7,452.6 |
| Father's occupation [b] | | | |
| At age 2 1/2 | 100 | 5.0 | 1.7 |
| At age 4 | 91 | 5.0 | 1.6 |
| Mother's education (years of formal schooling completed) | 100 | 10.9 | 2.8 |
| Father's education (years of formal schooling completed) | 100 | 11.7 | 3.4 |
| Child was only child ( 1 = no, 2 = yes ) | | | |
| At age 2 1/2 | 100 | ( 15% were only children ) | |
| At age 4 | 93 | ( 10% were only children ) | |
| Child's birth order ( 1 = first-born or only child, 2 = second born, etc. ) | 100 | 2.6 | 1.4 |
| Home Language [c] | | | |
| At age 2 1/2 | 75 | 20.3 | 32.0 |
| At age 4 | 70 | 18.4 | 29.8 |
| Performance | | | |
| MSCA General Cognitive Scale ( composite raw scores ) | | | |
| At age 2 1/2 | 67 | 23.4 | 12.0 |
| At age 3 | 69 | 48.7 | 18.4 |
| At age 3 1/2 | 74 | 62.8 | 20.3 |
| At age 4 | 70 | 84.0 | 23.0 |
| MSCA General Cognitive Index | | | |
| At age 2 1/2 | 67 | 93.2 | 14.3 |
| At age 3 | 69 | 98.6 | 12.7 |
| At age 3 1/2 | 74 | 97.9 | 12.6 |
| At age 4 | 70 | 94.9 | 13.9 |
| Caldwell Preschool Inventory | | | |
| At age 3 1/2 | 74 | 26.3 | 8.8 |
| At age 4 | 69 | 36.4 | 9.2 |
| Cattell Culture Fair Intelligence Test (mothers) | | | |
| Raw scores | 94 | 25.8 | 6.5 |
| Standardized scores | 94 | 88.8 | 15.0 |

*Note.* [a] Age in years.

[b] Father's occupation was measured using the following scale adapted from that used by the U.S. Bureau of the Census: 1 = private household workers, 2 = service workers ( except in private household ), 3 = laborers and farmers, 4 = equipment operators, 5 = craftsmen, foremen, and kindred workers, 6 = sales, clerical, and kindred workers, 7 = small business owners, or managers, or administrators, 8 = professional and technical, 9 = large business owners or managers.

[c] Mother speaks Spanish to child. Of all the verbalizations that the mother directs to the child in the course of everyday interactions, percentage that are in Spanish. ( One hundred minus this figure is the percentage in English and / or in a dialect that intermixes English and Spanish within sentences. )

BEST COPY AVAILABLE

attending school.) Each item belongs to one of the NYLS temperament categories: activity level, rhythmicity of biological functions, initial approach-withdrawal, adaptability, intensity of reaction, quality of mood, attention span and persistence, distractibility, and threshold of responsiveness. The respondent is asked to rate each item on a six-point scale (1 = almost never, 6 = almost always). The administrations were conducted orally. Although the oral administration was more costly and time consuming than the usual written administration, it was considered methodologically superior for this population because of the wide range in formal education and concomitant diversity in literacy levels.

## Performance Measures

Measures of the children´s general cognitive ability were obtained with the McCarthy Scales of Children´s Abilities (MSCA) (McCarthy, 1972). As a new alternative for users requiring a standardized test for young children, the MSCA is a well standardized instrument, carefully normed on a stratified, nationwide sample; efforts on the part of its developers to be sensitive to sociocultural diversity are evident in the design, standardization, and norming of this instrument. A most important consideration guiding the selection of this instrument, research has demonstrated the MSCA´s sound psychometric properties in a variety of ethnic samples, whether administered in English or Spanish. Such research also reveals that MSCA performance is a significant predictor of school achievement, again in a variety of ethnic samples, including Chicano and other Spanish-speaking children (for a review of research see Kaufman, 1982; see also Laosa, 1982, in press; Valencia, 1982, 1983). For the present sample, the Spearman-Brown split-half (odd-even items) reliability coefficients of the MSCA General Cognitive Scale were .94, .94, .91, and .94 at ages 2 1/2, 3, 3 1/2, and 4, respectively; the magnitudes of these coefficients were almost identical to those reported by McCarthy (1972) for children of the same age in the MSCA standardization sample. The General Cognitive Scale yields a composite raw score and a scaled ("index") score — the latter is a normalized, standardized score with M = 100 and SD = 16. The analyses were performed using both scores; they yielded practically identical results; the former are reported here.

Measures of the children's achievement were obtained with the Caldwell Preschool Inventory (CPSI). Developed particularly for use with economically disadvantaged Head Start children, the CPSI is intended to give a measure of achievement in areas regarded as necessary for success in school. It taps a range of verbal, quantitative, and perceptual-motor skills defined by teachers as expected of children in kindergarten (Cooperative Tests and Services, 1970b). Previous studies have yielded high reliability and validity indices for CPSI scores (Cooperative Tests and Services, 1970b; Gilbert & Shipman, 1972; Walker, Bane, & Bryk, 1973), including recent research on Chicano children (Laosa, 1982) and the normative investigations conducted in conjunction with the standardization of the CPSI Spanish language edition (Cooperative Tests and Services, 1974a). The most recent revision of the CPSI (Caldwell, 1970, 1974; Cooperative Tests and Services, 1970a, 1970b, 1974a, 1974b) was used in the present study. It contains 64 items and yields a total raw score based on the number of items answered correctly. For the present sample, the Kuder-Richardson Formula 20 reliability coefficients of the CPSI scores were .88 and .90 at ages 3 1/2 and 4, respectively; the magnitudes of these coefficients were very similar to those reported for both the English-speaking and the Spanish-speaking national standardization samples (Cooperative Tests and Services, 1970b, 1974a).

Measures of the mothers' nonverbal, fluid intelligence were obtained with the Cattell Culture Fair Intelligence Test (CFIT) (Institute for Personality and Ability Testing, 1973a, 1973b). The CFIT is designed to measure intelligence in a manner that reduces the influence of verbal fluency, cultural learning, and schooling level. Scale 2, Form A, was used in the present study. It contains 46 items and yields a raw score and a normalized, standardized score with $\underline{M} = 100$ and $\underline{SD} = 16$. The analyses were performed using both scores; they yielded practically identical results. Those based on the raw scores are reported here. Previous research in a variety of cultural and language groups indicates that the test yields adequately reliable and valid measurements of fluid intelligence (for reviews see Institute for Personality and Ability Testing, 1973a, 1973b). For the present sample, the Kuder-Richardson Formula 20 reliability coefficient of the CFIT scores was .82.

## Family Background Measures

The data on family background were obtained through structured interviews with the mothers. Each mother was interviewed twice; initially when the focal child was 2 1/2 years old and again when the child was 4. Some interview questions were administered only once, and others were repeated longitudinally, as Table 1 shows. Because the use of reports as a method of collecting data may be subject to response bias resulting from normative attitudes that may affect informants' judgments, special care was taken to minimize the possibility of such bias by employing and carefully training for the collection of the data only bilingual women who were indigenous to the ethnic, language, and geographic community from which they obtained the data.

## Administration of Measures

All the measures except the CFIT were administered individually in the participants' own homes; the CFIT was administered in small groups. The measures were administered by five trained examiners/interviewers; they were bilingual Chicano women from the same ethnolinguistic and geographic community as the participating families. Measurements on each family were taken initially when the child was 2 1/2 years old (plus or minus 2 months), and thereafter at 6-month intervals (plus or minus 2 weeks), with the last administration thus taking place within 2 months of the child's fourth birthday. To control for the child's chronological age with precision, strict adherence to the schedule was emphasized, and there were only a few slight deviations from it. At each longitudinal point, all the measures on a given family were taken after the data collector had visited the home several times to observe and establish rapport -- all of this taking place within approximately a week's time. As much as possible, all the data on a given family were collected by the same data collector. To minimize the possibility of expectancy effects, the examiners and families received no prior information regarding the specific research questions; they were informed, however, that the study sought to increase knowledge about the conditions surrounding children's development. Families were randomly assigned to data collectors.

The measures were administered in the participants' language or dialect (English, Spanish, or a dialect that switches between and/or mixes the two languages). For this purpose, the BSQ was professionally translated into local San Antonio Spanish; it was then backtranslated and the differences were resolved by discussion among the bilingual research staff. The same process was used for the other measures; for the CPSI and CFIT, the process built on their published Spanish versions. However, English was the preferred language of administration for the vast majority of the participants: To 84% of the mothers and 76% of the children (longitudinal averages), the measures were administered predominantly (i.e., at least 75% of the administration) in English.

## Staff Training and Monitoring

Considerable time and effort went into training the examiners/interviewers to collect reliable and valid data and into monitoring the collection, scoring, checking, and processing of the data in order to maintain the standards attained in training. For example, at random intervals throughout data collection, examiners/interviewers were paired together or with a field supervisor -- thus, while one examiner/interviewer administered the measures, another one, or the field supervisor, unobtrusively observed and independently marked a protocol; the two sets of protocols were later compared in supervisory meetings.

## Analytic Samples and Missing Data

There was very little longitudinal attrition. Of the 100 focal children who formed the sample at age 2 1/2 years, 94 were still available for data collection at age 4 years. BSQ data were available on 97 mothers when the children were 3 1/2 years old and again on all but three of the same mothers when their children were 4 years old. The analyses involving the MSCA, CPSI, and home language data are based on a randomly selected subsample of 75 children who were administered the MSCA at all four longitudinal points; all the other analyses are based on the entire sample. At each longitudinal point, very few data were missing on the available families; the instances of missing data were solved by deletion, as reflected in the sample sizes reported for each analysis.

## Results

### Interitem and Item-Category Correlations

A principal question regarding the internal structure of the data centered on whether the responses to the items composing a temperament category shared a common core; that is, whether for the sample at hand the individual items measured the same construct. Preliminary answers to this question were manifested in the means and standard deviations of the intercorrelations among the items composing each BSQ temperament category and in the correlations between the items and their respective total category scores.[4]

The mean interitem correlations tended to be low, ranging from .03 (threshold) to .14 (approach, mood) at age 3 1/2 and from .06 (threshold) to .14 (intensity, persistence) at age 4; in contrast, the standard deviations of the interitem correlations tended to be large -- ranging from .12 to .20 -- indicating a fairly wide spread about the mean interitem correlations. Accordingly, the median item-total correlations were low to moderate -- ranging from .07 (threshold) to .34 (approach) at age 3 1/2 and from .19 (threshold) to .36 (rhythmicity) at age 4 -- and there were considerable differences among the items composing any given category in the extent to which they correlated with the total category scores, including in most categories a few items with negative item-total correlations. A likely explanation of these results is that more than one dimension accounted for the variance in these Chicano mothers' responses to the items composing a category. The next analyses sought to test this hypothesis and to identify such dimensions.

## Dimensional Analyses

Requiring a procedure that would yield exact scores on the dimensions led to the selection of principal-components (with no interations) as the method of dimensional analysis. Principal-components analysis extracts from the unaltered intercorrelation matrix dimensions (components) that are expressed simply in terms of the observed variables, thus avoiding the approximation and estimation procedures that are required in factor analysis. To demonstrate that principal-components analysis is robust when the number of items is slightly larger than the number of cases, the analyses were performed twice, first with all 99 BSQ items and again with a reduced set. For the latter, an item was deleted from each of the temperament categories; selected for deletion from each category was the item with the lowest or a very low (or negative) corrected correlation with its total category score at both age levels, leaving a total of 90 items for analysis. The principal-components results obtained with slightly fewer items than cases were practically identical to those obtained with slightly more items than cases. Further, to determine the stability of solutions across methods of extracting dimensions, factor analyses were also performed, estimating the initial communalities by the absolute values of the maximum off-diagonal correlations. The factor-analytic solutions were very similar to those obtained by the principal-components analyses. Reported here are the results of the principal-components analyses based on the 99 items.

Components were extracted from the item intercorrelations, separately by age. Thirty-three components at age 3 1/2 and 32 at age 4 had eigenvalues greater than or equal to one; at each age, these components together accounted for 80% of the total variance in the responses (Appendix). To facilitate the interpretation of the componential structures, a series of rotations were performed. Initially, the number of components rotated at each age level was determined by the number of eigenvalues $\geq 1$, and the components were rotated to both orthogonal (varimax) and oblique (direct oblimin, delta = 0) solutions. The intercorrelations among the oblique components were either near zero or very low -- ranging from -.15 to .14 at age 3 1/2 and from -.18 to .15 at age 4 -- suggesting an absence of higher-order components. The different rotational techniques yielded similar solutions. The results of scree tests and a close scrutiny of the components in each solution suggested the presence of no more than five relatively broad, nontrivial, and interpretable dimensions in the responses.

A second set of rotations was therefore performed, this time orthogonally rotating only the first five components. To determine the stability of the componential structures over the 6-month interval, component scores were calculated and the cross-age product-moment correlations of the components were computed.[5] The five-component solutions revealed the presence of only two broad dimensions that were largely continuous and stable over this period. One dimension was reflected in Component I at age 3 1/2 and Component II at age 4. About half (43%) of the items defining Component I at age 3 1/2 also defined Component II at age 4. Additional evidence of continuity for this dimension came from an inspection of the items, which suggested that although not identical, the respective sets of items defining these two components sampled broadly overlapping domains of behavior. A moderate cross-age correlation between these two components provided evidence of stability ($r$ = .42, $p < .001$, two-tailed test). The other dimension was reflected in Component II at age 3 1/2 and Component I at age 4. The majority (63%) of the items defining Component II at age 3 1/2 also defined Component I at age 4; moreover, the respective items defining these two components, although not identical, appeared to represent broadly overlapping domains of behavior. The stability of this dimension was reflected in a substantial cross-age correlation between components ($r$ = .50, $p < .001$, two-tailed test).

The remaining components of the five-component solutions were difficult to interpret on the basis of the items that defined them and generally lacked continuity and stability. The exception was Component V at age 4, which was labeled "reaction to food"; five of the eight items that defined it focused on the child's reactions in the feeding situation. No such component emerged in the corresponding five-component solution at the younger age, although the preceeding solutions involving rotations of all the components with eigenvalues $\geq$ 1 did yield a component reflecting sensitivity to food at each age. Because of the specific nature of this component, however, its value as a general dimension of temperament can be questioned. Similar dimensions reflecting the child's reactions to food have been found in U. S. middle-class samples and in Swedish samples (Bohlin, Hagekull, & Lindhagen, 1981; Hagekull, Lindhagen, & Bohlin, 1980; Huitt & Ashton, 1982; Rowe & Plomin, 1977). Aside from the feeding dimension, neither the patterns of intercorrelations among the items, nor the dimensions extracted therefrom, revealed clusters that could be unambiguously identified in the present BSQ data as dimensions reflecting particular types of situations or specific response modalities.

A final set of rotations was therefore performed, this time orthogonally rotating only the first two components. The cross-age correlations between the resulting solutions and the distributions of items by components and BSQ categories are shown, respectively, in Tables 2 and 3.

The two components were largely continuous and stable over the 6-month span between 3 1/2 and 4 years of age. Evidence of continuity came from the cross-age overlap in the sets of items defining the components. Eighty percent of the items that defined Component I at age 3 1/2 also defined it at age 4; conversely, 69% of the items that defined this component at age 4 also defined it at age 3 1/2. Sixty-eight percent of the items that defined Component II at age 3 1/2 also defined it at age 4; conversely, 62% of the

Table 2

Cross-Age Correlations of Components

| | Age 4 | |
|---|---|---|
| | Component I | Component II |
| Age 3 1/2 | | |
| Component I | .68* | .11 |
| Component II | .02 | .60* |

Note. N = 94.

* $p < .001$, two-tailed tests.

Table 3

Classification of Items by Components and BSQ Categories, by Age

| BSQ category | Item no. | Component | | BSQ category | Item no. | Component | | BSQ category | Item no. | Component | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Age 3½ | Age 4 | | | Age 3½ | Age 4 | | | Age 3½ | Age 4 |
| Activity | 4. | I | I | Adaptability | 3. | II | I | Persistence | 27. | -- | II |
| | 6. | -- | -- | | 8. | I | -- | | 33. | -- | I |
| | 9. | -- | -- | | 10. | -- | -- | | 35. | II | II |
| | 13. | II | II | | 15. | I | I | | 39. | I | -- |
| | 14. | I | I | | 19. | -- | -- | | 40. | II | II |
| | 24. | -- | II | | 28. | II | II | | 71. | I | I |
| | 26. | I | I | | 55. | I | -- | | 73. | -- | -- |
| | 32. | I | I | | 56. | -- | I | | 83. | I | II |
| | 44. | II | II | | 61. | I | II | | 90. | -- | -- |
| | 58. | II | I | | 63. | -- | -- | | 93. | II | II |
| | 70. | I | I | | 65. | I | I | Distractibility | 2. | -- | I |
| | 87. | -- | I | | 80. | -- | -- | | 17. | I | I |
| | 94. | I | I | Intensity | 7. | -- | I | | 48. | -- | I |
| Rhythmicity | 11. | II | II | | 20. | -- | II | | 51. | II | -- |
| | 23. | II | II | | 30. | II | II | | 66. | II | II |
| | 36. | II | II | | 41. | I | I | | 77. | II | -- |
| | 47. | II | II | | 42. | -- | II | | 81. | II | -- |
| | 49. | II | II | | 45. | II | II | | 85. | II | I |
| | 52. | I | I | | 46. | II | II | | 89. | -- | -- |
| | 62. | II | II | | 53. | I | I | | 95. | -- | I |
| | 75. | II | I | | 76. | -- | II | Threshold | 16. | II | II |
| | 84. | -- | -- | | 82. | I | I | | 18. | -- | -- |
| Approach | 12. | -- | II | | 92. | II | II | | 22. | II | II |
| | 25. | -- | II | | 99. | II | -- | | 37. | -- | II |
| | 31. | II | II | Mood | 1. | I | I | | 57. | II | II |
| | 43. | -- | II | | 5. | -- | II | | 59. | -- | -- |
| | 50. | -- | -- | | 29. | -- | II | | 60. | II | II |
| | 54. | II | -- | | 34. | I | I | | 74. | II | II |
| | 67. | II | -- | | 38. | -- | II | | 79. | -- | -- |
| | 68. | I | -- | | 64. | I | I | | 88. | I | I |
| | 86. | -- | -- | | 69. | I | I | | 91. | II | -- |
| | 98. | -- | -- | | 72. | I | I | | | | |
| | | | | | 78. | I | I | | | | |
| | | | | | 96. | I | I | | | | |
| | | | | | 97. | I | I | | | | |
| | | | | | 100. | I | I | | | | |

*Note*. Dashes indicate loadings with absolute values < .30 on both components for a given age.

BEST COPY AVAILABLE

items that defined this component at age 4 also defined it at age 3 1/2. Additional evidence of continuity can be adduced from the fact that, for each component, the items not shared in common over age periods nevertheless appeared to represent the same broad domain of behavior as did the items shared in common.

The differences across time in the patterns of items defining the dimensions may reflect what Emmerich (1964, 1966) has generically termed minor developmental transformations (see also Kagan, 1971; Moss & Susman, 1980; Wohlwill, 1980). For example, the fact that item 3 ("child can be coaxed out of a forbidden activity") shifted from Component II at age 3 1/2 to Component I at age 4 may indicate that, at least in the mothers' cognitive representations, this aspect of the child's behavior underwent a transformation in meaning. Several other items were observed to shift dimensions. In addition, a few items did not load saliently on either component at the younger age, but did so at the older age, thus seeming to "emerge" as behavioral indices of these dimensions in the course of development between 3 1/2 and 4 years of age. Conversely, a few items exhibited salient loadings at the earlier age, but failed to do so at the older age -- thus seeming to cease functioning as defining features of these dimensions. Although it is interesting to speculate about the meaning of each apparent transformation, the probability that for any given item the difference might have occurred by chance is unknown.

Evidence of stability came from the substantial cross-age correlations of the component scores: $\underline{r} = .68$ for Component I and $\underline{r} = .60$ for Component II.

It is worth noting that the stability of the component scores was greater than that of the total scores on any BSQ temperament category. The cross-age correlations of the BSQ category scores ranged from .20 (mood) to .59 (activity), median $\underline{r}$ = 51.

As expected on the basis of the preceding analyses, the magnitudes of the final communalities in the two-component solutions were generally low, a reminder that there remained much unaccounted variance; nevertheless, these were the optimal solutions that met the criteria of short-term continuity and stability.

It is evident from this series of analyses that, at least for the present sample of Chicano mothers, the BSQ temperament categories did not possess componential unity; the principal-components analyses thus corroborated the hypothesis suggested by the interitem and item-total correlations. The distributions of items by components revealed, however, a fair degree of componential integrity for certain clusters of categories, as Table 3 shows. Of the items with salient loadings on either component, the vast majority of those belonging to the mood, activity, and adaptability categories, respectively, loaded on Component I, and the vast majority of those belonging to the rhythmicity, threshold, approach, and intensity categories, respectively, loaded on Component II. In contrast, the persistence and distractibility items were distributed in about equal parts between the two components.

Labels for the two components were suggested by the items defining them. Component I was labeled mood/manageability, and Component II, rhythmicity/responsivity. High mood/manageability scores identified mothers who rated their children as preponderantly negative in mood, crying frequently, fussy, slow in adapting to changes in the environment, and resistant to parental discipline and control. High rhythmicity/ responsivity scores characterized mothers who rated their children as irregular in biological functions, frequently unresponsive to sensory and social stimuli, often withdrawing from new stimuli and situations, low in emotional expressiveness, and nonpersistent.

## Sex Comparisons

Boys and girls were practically identical in the means, variances, and cross-age correlations of the component scores.

## Correlations and Multiple Regressions

Correlations and multiple regressions were calculated between the component scores of the two-component solutions and the variables listed in Table 1. Pearson product-moment and Spearman rank-order (corrected for ties) correlations yielded very similar results; the former are reported here. A visual inspection of the bivariate scatter plots revealed no unambiguous evidence of curvilinearity.

The children's performance on the tests of ability and achievement correlated significantly with Component I scores, both concurrently and predictively -- as Table 4 shows. Among the predictive coefficients, it is

Table 4

Predictive and Concurrent Correlations of the Children's Performance on the Cognitive Ability and Achievement Tests with the Components of Temperament Perceptions

| Test and Age | Component | | | |
|---|---|---|---|---|
| | Age 3 1/2 | | Age 4 | |
| | I | II | I | II |
| MSCA General Cognitive Scale | | | | |
| Age 2 1/2 | -.29* | .09 | -.15 | .15 |
| Age 3 | -.22 | .19 | -.21 | .19 |
| Age 3 1/2 | -.24* | .05 | -.20 | .04 |
| Age 4 | -.20 | .16 | -.24* | .12 |
| Caldwell Preschool Inventory | | | | |
| Age 3 1/2 | -.30** | .14 | -.22 | .08 |
| Age 4 | -.31** | .05 | -.26* | -.01 |

*Note*. Age in years. *N* = 64-74. Multiple *R* = .33 when the Caldwell Preschool Inventory scores at age 4 were regressed on the four components. Multiple *R* = .29 when the MSCA General Cognitive Scale scores at age 4 were regressed on the four components.

$^{*}p < .05$, $^{**}p < .01$; two-tailed tests.

particularly remarkable that the children's ability scores at age 2 1/2 years significantly predicted their temperament as viewed by their mothers a year later. In contrast, the mothers' performance on the Cattell Culture Fair Intelligence Test did not correlate significantly with the component scores ($N = 94$, $p > .05$, two-tailed tests).

Of the other variables, only home language correlated significantly with the component scores. Very low and near zero correlations were obtained between the measures of socioeconomic status and parents' education, on the one hand, and the component scores, on the other -- only one of the 24 correlations was signficant ($N = 91\text{-}97$, $p < .05$, two tailed tests). Similarly, near-zero and nonsignificant correlations were obtained between the measures of sibling structure and the component scores ($N = 93\text{-}97$, $p > .05$, two-tailed tests). In contrast, home language correlated significantly with Component II scores, as Table 5 shows.

## Discussion

The mood/manageability dimension obtained in the present study of Chicano families appears closely related to dimensions identified in previous research on other ethnic groups. Analyses of temperament ratings in Swedish samples have yielded a *manageability* dimension, which represented the parents' appraisals of their children's manageability and the children's mood quality (Bohlin et al., 1981; Hagekull et al., 1980). The present mood/manageability dimension in the Swedish samples seems related to the *fussiness* and *irritability* dimensions identified in several North American studies (e.g., Bates, Freeland, & Lounsbury, 1979; Birns, Barten, & Bridger, 1969; Goldsmith

Table 5

Predictive and Concurrent Correlations Between Home Language and Components of Temperament Perceptions

| Home Language | Component |  |  |  |
|---|---|---|---|---|
|  | Age 3 1/2 |  | Age 4 |  |
|  | I | II | I | II |
| Mother speaks Spanish to child |  |  |  |  |
| At age 2 1/2 | .10 | -.44** | .07 | -.31* |
| At age 4 | .09 | -.36* | .00 | -.31* |

Note. Age in years. N = 70-74.

* $p <$ .01, ** $p <$ .001; two-tailed tests.

& Gottesman, 1981). It may be related also to Buss and Plomin's (1975) emotionality dimension. Further, some elements of the Chicano mood/manageability dimension are reminiscent of the permissiveness versus restrictiveness and control versus autonomy dimensions found in the circumplex ordering of maternal ratings in studies of childrearing practices (e.g., Schaefer, 1959; for a review see Martin, 1975). Such commonalities are relevant to the testing of hypotheses about the cross-cultural universality of temperament perceptions.

Similarly, the mood/manageability dimension shares some elements in common with the construct of difficult temperament advanced by the NYLS group (Thomas et al., 1968). There is, however, a notable difference. In the NYLS conceptual framework, mood and rhythmicity are grouped together as elements of the same conceptual dimension, whereas in the present data these two elements belong to separate and independent dimensions. Ironically, in this regard the present data are quite consistent with the NYLS data, although inconsistent with the NYLS conceptual framework. In the NYLS data, as in the present data, mood and rhythmicity loaded on separate and independent dimensions (see Thomas et al., 1968). Nevertheless, the NYLS team retained rhythmicity in their definition of difficult temperament, arguing that "irregularity does make for greater difficulty in the attainment of regular sleep and feeding schedules and toilet training in the early childhood years, and appears functionally related, even if not psychometrically, to the other four categories of difficult temperament" (Thomas et al., 1982, p. 5). From the standpoint of population validity, this argument is particularly difficult to reconcile with another facet of the NYLS data, namely, the comparison between Jewish professional and Puerto Rican low-income families (Thomas et al., 1974), which indicated that the functional relationship between rhythmicity and difficultness is culture-specific and not generalizable across sociocultural groups. These inconsistencies raise questions about the construct validity of at least one aspect of the NYLS conceptual definition of difficult temperament.

Previous studies of maternal perceptions of temperament have extracted at least three and often four or more dimensions. Thomas and his coworkers (1968) extracted three; Buss and Plomin (1975), Garside, Birch, Scott, Chambers, Kolvin, Tweddle, and Barber (1975), McDevitt (1976), and Bates et al. (1979), four; Huitt and Ashton (1982) and Lerner, Palermo, Spiro, and Nesselroade (1982), five; Bohlin et al. (1981), seven; and Hagekull et al. (1980), eight. Why fewer dimensions in the present study? Unlike previous studies, the selection of dimensions in the present study was determined by the two-fold criterion of short-term longitudinal continuity and stability; only two broad dimensions thus qualified for retention. The importance of this criterion lies in the fact that it is possible for a dimension to be extracted because of a chance patterning of intercorrelations; this possibility is considerably reduced by retaining only the structures that are repeatable (i.e., reliable). Obversely, the risk in applying this criterion is that real but unstable or ephemeral dimensions, if they exist, will be discarded. A challenge for future research is to develop methodologies for reliably distinguishing unstable and ephemeral dimensions from those that are merely the result of error, meaningless residual variance, or chance patternings in the data.

The present study found a relationship between maternal perceptions of temperament and the child's cognitive performance. Children who scored lower on tests of cognitive ability and achievement were characterized by their mothers as being more negative in mood, crying more frequently, fussier, slower in adapting to changes in the environment, and more resistant to parental discipline and control. This finding appears consistent with previous research on other sociocultural populations (for reviews see Campos et al., 1983; Lerner & Lerner, 1983). Needed now are studies to elucidate the mechanisms that account for this relationship.

The analyses also uncovered a relationship between home language and temperament perceptions. A more frequent use of Spanish for mother-child interaction in the home -- and conversely, a less frequent use of English or of the admixed English-Spanish dialect -- was associated with maternal perceptions of the child as more rhythmic in biological functions, more responsive to sensory and social stimuli, more frequently approaching new stimuli and situations, emotionally more expressive, and more persistent. This finding lends credence to the hypothesis that perceptions of temperament -- if not temperament itself -- are subject to sociocultural influences, insofar as the maintenance or loss of Spanish in Chicano households indicates degree of acculturation.

## Acknowledgment

The research reported in this article was supported in part by Grant 90-C-1257 (Luis M. Laosa, Principal Investigator) from the United States Administration for Children, Youth, and Families. Portions of the data were presented at the Fourth Occasional Temperament Conference, Salem, Massachusetts, October 1982 and at the Annual Meeting of the American Educational Research Association, New Orleans, Louisiana, April 1984.

The author gratefully acknowledges the assistance of Carolina Gomez in the coordination of data collection; of Carolina Gomez, Anna Martinez, Irma Pena, and Linda Ramirez-Ottis in the collection of data; of Debra Friedman, Bernardo Ferdman, Margaret Sullivan, Margaret Redman, and Connie Struve in the scoring and coding of protocols; of William Nemceff in the preparation of computerized files; and of Robin Roth in typing the manuscript and operating the word processor.

# References

Bates, J. E. (1980). The concept of difficult temperament. *Merrill-Palmer Quarterly*, *26*, 299-319.

Bates, J. E., Freeland, C. A. B., & Lounsbury, M. L. (1979). Measurement of infant difficultness. *Child Development*, *50*, 794-803.

Birns, B., Barten, S., & Bridger, W. H. (1969). Individual differences in temperamental characteristics of infants. *Transactions of the New York Academy of Sciences*, *31*, 1071-1082.

Bohlin, G., Hagekull, B., & Lindhagen, K. (1981). Dimensions of infant behavior. *Infant Behavior and Development*, *4*, 83-96.

Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality development*. New York: Wiley.

Caldwell, B. M. (1970). *Cooperative Preschool Inventory: Revised edition -- 1970*. Princeton, NJ: Educational Testing Service.

Caldwell, B. M. (1974). *Cooperative Preschool Inventory: Spanish edition*. Princeton, NJ: Educational Testing Service.

Campos, J. J., Barrett, K. C., Lamb, M. E., Goldsmith, H. H., & Stenberg, C. (1983). Socioemotional development. In M. M. Haith & J. J. Campos (Eds.), *Infancy and developmental psychobiology* [vol. 2 of P. H. Mussen (Ed.), *Handbook of child psychology*]. New York: Wiley.

Cooperative Tests and Services. (1970a). *Preschool Inventory revised edition -- 1970: Directions for administering and scoring*. Princeton, NJ: Educational Testing Service.

Cooperative Tests and Services. (1970b). *Preschool Inventory revised edition -- 1970: Handbook*. Princeton, NJ: Educational Testing Service.

Cooperative Tests and Services. (1974a). *Cooperative Preschool Inventory, Spanish edition, preliminary edition: Handbook*. Princeton, NJ: Educational Testing Service.

Cooperative Tests and Services. (1974b). *Preschool Inventory, Spanish edition, preliminary edition: Directions for administering and scoring.* Princeton, NJ: Educational Testing Service.

Emmerich,W. (1964). Continuity and stability in early social development. *Child Development*, *35*, 311-332.

Emmerich, W. (1966). Continuity and stability in early social development: II. Teacher ratings. *Child Development*, *37*, 17-27.

Emmerich, W. (1968). Personality development and concepts of structure. *Child Development*, *39*, 671-690.

Garside, R. F., Birch, H., Scott, D. McI., Chambers, S., Kolvin, I., Tweddle, E. G., & Barber, L. M. (1975). Dimensions of temperament in infant school children. *Journal of Child Psychology and Psychiatry*, *16*, 219-231.

Gilbert, L. E., & Shipman, V.C. (1972). Preschool Inventory. In V. C. Shipman (Ed.), *Disadvantaged children and their first school experiences: Technical report series* (PR 72-27). Princeton, NJ: Educational Testing Service.

Goldsmith, H. H., & Gottesman, I. I. (1981). Origins of variation in behavioral style: A longitudinal study of temperament in young twins. *Child Development*, *52*, 91-103.

Hagekull, B., Lindhagen, K., & Bohlin, G. (1980). Behavioral dimensions in one-year- olds and dimensional stability in infancy. *International Journal of Behavioral Development*, *3*, 351-364.

Hubert, N. C., Wachs, T. D., Peters-Martin, P., & Gandour, M. J. (1982). The study of early temperament: Measurement and conceptual issues. *Child Development*, *53*, 571-600.

Huitt, W. G., & Ashton, P. T. (1982). Parents' perception of infant temperament: A psychometric study. *Merrill-Palmer Quarterly*, *28*, 95-109.

Institute for Personality and Ability Testing (1973a). *Measuring intelligence with the Culture Fair Tests: Manual Scales 2 and 3.* Champaign, IL: Institute for Personality and Ability Testing.

Institute for Personality and Ability Testing (1973b). *Technical Supplement for the Culture Fair Intelligence Tests: Scales 2 and 3*. Champaign, IL: Institute for Personality and Ability Testing.

Kagan, J. (1971). *Change and continuity in infancy*. New York: Wiley.

Kaufman, A. S. (1982). An integrated review of almost a decade of research on the McCarthy Scales. In T. R. Kratochwill (Ed.), *Advances in school psychology* (Vol. 2). Hillsdale, NJ: Erlbaum.

Keogh, B. K., & Pullis, M. E. (1980). Temperament influences on the development of exceptional children. In B. K. Keogh (Ed.), *Advances in special education (Vol. 1): Basic constructs and theoretical orientations*. Greenwich, CT: JAI Press.

Laosa, L. M. (1977). Socialization, education, and continuity: The importance of the sociocultural context. *Young Children*, *32*, 21-27.

Laosa, L. M. (1979). Social competence in childhood: Toward a developmental, socioculturally relativistic paradigm. In M. W. Kent & J. E. Rolf (Eds.), *Primary prevention of psychopathology (Vol. 3): Social competence in children*. Hanover, NH: University Press of New England.

Laosa, L. M. (1981a). Maternal behavior: Sociocultural diversity in modes of family interaction. In R. W. Henderson (Ed.), *Parent-child interaction: Theory, research, and prospects*. New York: Academic Press.

Laosa, L. M. (1981b, October). *Statistical explorations of the structural organization of maternal teaching behaviors in Chicano and non-Hispanic White families*. Invited paper presented at the Conference on the Influences of Home Environments on School Achievement, Wisconsin Research and Development Center for Individualized Schooling, School of Education, University of Wisconsin, Madison.

Laosa, L. M. (1982). Psychometric characteristics of Chicano and non-Hispanic White children's performance on the Preschool Inventory. *Journal of Applied Developmental Psychology*, *3*, 217-245.

Laosa, L. M. (in press). Ethnic, socioeconomic, and home language influences upon early performance on measures of abilities. *Journal of Educational Psychology*.

Lerner, R. M., & Lerner, J. V. (1983). Temperament-intelligence reciprocities in early childhood: A contextual model. In M. Lewis (Ed.), *Origins of intelligence: Infancy and early childhood* (2nd ed.). New York: Plenum.

Lerner, R. M., Palermo, M., Spiro, A., III, & Nesselroade, J. R. (1982). Assessing the dimensions of temperamental individuality across the life span: The Dimensions of Temperament Survey (DOTS). *Child Development*, *53*, 149-159.

Martin, B. (1975). Parent-child relations. In F. D. Horowitz, E. M. Hetherington, S. Scarr-Salapatek, & G. M. Siegel (Eds.), *Review of child development research* (Vol. 4). Chicago: University of Chicago Press.

McCarthy, D. (1972). *Manual for the McCarthy Scales of Children's Abilities*. New York: The Psychological Corporation.

McDevitt, S. C. (1976). *A longitudinal assessment of continuity and stability in temperamental characteristics from infancy to early childhood*. Unpublished Ph. D. dissertation, Temple University.

McDevitt, S. C., & Carey, W. B. (1975). *Behavioral Style Questionnaire*. Available from S. C. McDevitt, Devereux Center, 6436 East Sweetwater, Scottsdale, Arizona 85254.

McDevitt, S. C., & Carey, W. B. (1978). The measurement of temperament in 3-7 year old children. *Journal of Child Psychology and Psychiatry*, *19*, 245-253.

Moss, H. A., & Susman, E. J. (1980). Longitudinal study of personality development. In O. G. Brim, Jr. & J. Kagan (Eds.), *Constancy and change in human development*. Cambridge, MA: Harvard University Press.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. (1975). *SPSS: Statistical package for the social sciences* (2nd ed.). New York: McGraw-Hill.

Rowe, D. C., & Plomin, R. (1977). Temperament in early childhood. *Journal of Personality Assessment*, *41*, 150-156.

Rutter, M. (1982). Temperament: Concepts, issues, and problems. In R. Porter & G. M. Colling (Eds.), *Temperamental differences in infants and young children* (Ciba Foundation symposium 89). London: Pitman.

Schaefer, E. S. (1959). A circumplex model for maternal behavior. Journal of Abnormal and Social Psychology, 59, 226-235.

Thomas, A., & Chess, S. (1977). Temperament and development. New York: Brunner/Mazel.

Thomas, A., Chess, S., & Birch, H. G. (1968). Temperament and behavior disorders in children. New York: New York University Press.

Thomas, A., Chess, S., Birch, H. G., Hertzig, M. E., & Korn, S. (1963). Behavioral individuality in early childhood. New York: New York University Press.

Thomas, A., Chess, S., & Korn, S. J. (1982). The reality of difficult temperament. Merrill-Palmer Quarterly, 28, 1-20.

Thomas, A., Chess, S., Sillen, J., & Mendez, O. (1974). Cross-cultural study of behavior in children with special vulnerabilities to stress. In D. F. Ricks, A. Thomas, & M. Roff (Eds.). Life history research in psychopathology (Vol. 3). Minneapolis: University of Minnesota Press.

U. S. Bureau of the Census. (1979). Persons of Spanish origin in the United States: March 1978 (Current Population Reports, Series P-20, No. 339). Washington, DC: U. S. Government Printing Office.

U. S. Bureau of the Census. (1982). 1980 census of population. Persons of Spanish origin by state: 1980. (Supplementary Report PC80-S1-7). Washington, DC: U. S. Government Printing Office.

Valencia, R. R. (1982). Predicting academic achievement of Mexican American children: Preliminary analyses of the McCarthy Scales. Educational and Psychological Measurement, 42, 1269-1278.

Valencia, R. R. (1983). Stability of the McCarthy Scales of Children's Abilities over a one-year period for Mexican American children. Psychology in the Schools, 20, 29-34.

Walker, D. R., Bane, M. J., & Bryk, T. (1973, August). The quality of the Head Start planned variation data (Vols. 1 & 2). Cambridge, MA: Huron Institute.

Wohlwill, J. F. (1980). Cognitive development in childhood. In O. G. Brim, Jr. & J. Kagan (Eds.), *Constancy and change in human development*. Cambridge, MA: Harvard University Press.

# Appendix

Eigenvalues and Percentages of Variance for Components with Eigenvalues ≥ 1, by Age

| Component | Age 3 1/2 |  | Age 4 |  |
|---|---|---|---|---|
|  | Eigenvalue | % variance | Eigenvalue | % variance |
| 1 | 8.38 | 8.5 | 9.25 | 9.3 |
| 2 | 6.30 | 6.4 | 7.58 | 7.7 |
| 3 | 4.23 | 4.3 | 4.94 | 5.0 |
| 4 | 3.72 | 3.8 | 4.08 | 4.1 |
| 5 | 3.53 | 3.6 | 3.51 | 3.5 |
| 6 | 3.24 | 3.3 | 3.26 | 3.3 |
| 7 | 3.12 | 3.1 | 3.07 | 3.1 |
| 8 | 3.03 | 3.1 | 2.92 | 3.0 |
| 9 | 2.88 | 2.9 | 2.74 | 2.8 |
| 10 | 2.59 | 2.6 | 2.57 | 2.6 |
| 11 | 2.56 | 2.6 | 2.41 | 2.4 |
| 12 | 2.49 | 2.5 | 2.27 | 2.3 |
| 13 | 2.29 | 2.3 | 2.17 | 2.2 |
| 14 | 2.16 | 2.2 | 2.12 | 2.1 |
| 15 | 2.03 | 2.0 | 2.10 | 2.1 |
| 16 | 1.97 | 2.0 | 1.91 | 1.9 |
| 17 | 1.91 | 1.9 | 1.78 | 1.8 |
| 18 | 1.84 | 1.9 | 1.77 | 1.8 |
| 19 | 1.74 | 1.8 | 1.74 | 1.8 |
| 20 | 1.73 | 1.7 | 1.66 | 1.7 |
| 21 | 1.68 | 1.7 | 1.59 | 1.6 |
| 22 | 1.59 | 1.6 | 1.55 | 1.6 |
| 23 | 1.54 | 1.6 | 1.46 | 1.5 |
| 24 | 1.52 | 1.5 | 1.43 | 1.4 |
| 25 | 1.46 | 1.5 | 1.42 | 1.4 |
| 26 | 1.41 | 1.4 | 1.33 | 1.3 |
| 27 | 1.33 | 1.3 | 1.32 | 1.3 |
| 28 | 1.30 | 1.3 | 1.23 | 1.2 |
| 29 | 1.26 | 1.3 | 1.14 | 1.2 |
| 30 | 1.23 | 1.2 | 1.10 | 1.1 |
| 31 | 1.11 | 1.1 | 1.02 | 1.0 |
| 32 | 1.06 | 1.1 | 1.01 | 1.0 |
| 33 | 1.03 | 1.0 |  |  |

# Footnotes

[1] I use the term "Chicano" or "Mexican American" to refer to (a) persons who trace their lineage to Indo-Hispanic forebears who resided within Spanish or Mexican territory that is not part of the southwestern United States, (b) persons with cultural and linguistic roots in the Spanish or mixed Spanish and Indian history of this region, (c) persons whose parents or more remote ancestors immigrated to the United States from Mexico, and (d) persons who were born in Mexico and now hold U. S. citizenship or otherwise reside in the United States. "Persons of Mexican origin" is the term used by the U. S. Bureau of the Census, which in 1980 counted over 8.7 million such persons in the United States (U. S. Bureau of the Census, 1982).

[2] The only exceptions were two fathers -- one was of mixed Mexican and Italian parentage and the other was Puerto Rican.

[3] The only exceptions were three mothers, who had part-time employment.

[4] For purposes of the analyses, the scale values of certain items were reversed, such that for every item a high rating would indicate a high score on its category. The item-total correlations were corrected for the spuriousness that can result from the overlap between the items and their respective category scores; this correction involved calculating the correlation between an item and the category score based upon the other remaining $\underline{n} - 1$ items in the category of $\underline{n}$ items.

[5] Exact component-score coefficients were calculated from the rotated pattern matrix. For five respondents, who had missing data on few items, component scores were calculated as weighted products of the existing data (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975).

# TEMPERAMENT AND BEHAVIOR DISORDERS: TEACHERS' PERCEPTIONS OF STUDENT CONTROL

Michael Pullis

University of Texas at Dallas
School of Human Development

Temperament or behavioral style refers to the typical way an individual responds to or interacts with his or her environment. Thomas and Chess (1977) define temperament as the "how" of behavior and contrast it with actual behavior (what children do) or motivation (the why of behavior). Early in their longitudinal study of children's development, they identified nine descriptive dimensions of temperament: activity level, attention span and persistence, distractibility, approach/withdrawal, adaptability, mood, threshold of responding, intensity of responding, and rhythmicity. These dimensions were used to create a rating form by which parents characterized their child's usual way of responding to objects, people, and events in their environment. Further, Thomas, Chess, and Birch (1968) demonstrated that a mismatch between certain of these traits and environmental factors (primarily parenting strategies) can lead to behavioral adjustment difficulties. These problematic mismatches were most likely developed when the child was characterized as "Difficult," evidencing extremely negative tendencies in the behavioral dimensions of rhythmicity, withdrawal, adaptability, intensity of behavior, and negative mood.

While Difficult children appeared to be at a higher risk for developing problems, a number of children with Easy temperament profiles also were found to adjust poorly. Thomas et al. maintained that in any given case difficulties emerged as the result of a "lack of fit" between the child's behavioral tendencies and their parents' strategies and/or attitudes. Clinical intervention with these families consisted of assessing the child's temperament and providing parents with strategies that would accommodate to rather than conflict with their child's characteristics, focusing on creating a more productive match as compared to attempting to modify (often through unpleasant discipline) the overt behavior of the child.

Temperament continues to be a topic of research appearing often in literature concerned with stability of behavior across development, prediction of adjustment difficulties or risk, and adult-child interactions. Relatively little research has appeared that examines the role of children's temperament characteristics in their adjustment to and performance in school settings. Some studies have cited modest but statistically reliable correlations between parent ratings of temperament and various measures related to school adjustment such as achievement test scores and teacher ratings. However, Thomas and Chess suggest that within the school environment, teachers might be the most appropriate source of information about relevant student characteristics. For this reason, they developed their Teacher Temperament Questionnaire (TTQ) to describe students' temperament within various school interactions. Keogh, Pullis, and Cadwell (1982) have reported that this instrument can be modified to a reduced format that contains three independent factors of student temperament -- task orientation, adaptability, and reactivity. These three factors contain items from all of the original dimensions of temperament, thus retaining the integrity of the Thomas and Chess approach, while yielding a psychometrically sound research instrument.

Research studies that have used this short form of the TTQ have shown that this instrument is valid for typical elementary students (Pullis & Cadwell, 1982), children educationally at risk (Pullis & Cadwell, in press), and learning disabled students (Pullis, in press). Furthermore, these studies have consistently shown that these temperament factors significantly influence teachers' classroom decision-making policies. Even when considered simultaneously with more traditional measures of achievement and intelligence, temperament variables have been linked with a variety of teacher decisions -- monitoring needs, behavior management approaches, instructional modifications, etc.

Very few studies have examined the relationship between emotionally disturbed (ED) students' temperament and their progress in school-based intervention settings. Those that have, continued to use parent ratings as the source of information about temperament and have tended to limit their discussion to Difficult children. Students categorized as ED are often diagnosed with labels such as psychotic or conduct disordered often emanating from psychological or psychiatric classification schemes. Kauffman (1981) and others have criticized these labels as having little utility with respect to intervention approaches and have called for more relevant descriptions of student behavior. Given the findings that implicate temperament characteristics as important influences on teachers' decision-making in classroom situations, it might be useful to delineate patterns of temperament factors that may be differentially distributed across different subgroups of ED students.

The present study was designed to examine these differential patterns as well as to determine if and how these characteristics were related to ED teachers´ classroom decision-making and the stress they reported as part of their occupation. Of particular interest were the findings concerning the teachers´ perceptions of how much control students had over the problematic behavior they exhibited. This notion of perceived control is an important clinical issue in the treatment or intervention of childhood behavior disorders. Kauffman (1981) and others have suggested that when teachers believe that students have control over their behavior there is a tendency for the teachers to develop more negative affective responses to that child. Because they believe that the child could control their behavior, continued episodes of disruptive or nonproductive student behavior mean that the child is not putting forth effort with respect to improving their behavior. Thus, teachers often report feelings of anger, disappointment or frustration with this type of student. These negative affective responses can lead to more punitive methods of control as well as teacher-student relationships marked by power struggles and sarcasm (Brophy & Rohrkemper, 1981).

The notion of perceived control is also related to the clinical work of the Thomas and Chess group. They report that a key element of their clinical intervention with their psychiatric treatment groups focused on helping parents understand that their child´s behavior reflected individual differences that were to some degree a function of congenital factors. They attempted to help the parents formulate a new interpretation of their child´s behavior away from the idea that the child's behavior was totally volitional or that they had been ineffective parents and somehow caused the child's problems. Thus, like Kauffman, they suggest that an important aspect of adult-child relationships lies in the understanding and interpretation of the child's behavior by the adult. In both cases, then, the goal is to increase objectivity through a more thorough understanding of the causes of children's problematic behavior. This objectivity can lead to two important changes in the adult-child relationship. First, negative affective responses can be reduced when the adult is willing to entertain the idea that the child may not have total control over their behavior. Increased objectivity then can lead to more sensitivity and motivation on the part of the adult and an improvement in the qualitative aspects of the relationship. Interactions with behavior disordered children can be extremely "emotionally charged" and objectivity can allow the adult to create some degree of distance from these problematic interactions and therefore not maintain continuous conflictual interchanges with the child.

The second advantage to adults' reconceptualizing their interpretation of children's behavior concerns discipline or management strategies. A focus of the clinical work of Thomas and Chess emphasized the creation of a more productive fit between the child's behavioral tendencies or style and the demands of the social and learning environment. They encouraged parents to have a clear understanding of their child's behavioral style, take these behavioral tendencies into account when planning family activities (accommodation), and respond to the child's problematic behavior with firmness and consistency. Their suggestions to parents parallel strategies that can be useful for teachers through more formalized procedures such as assessment, proactive planning of learning and social activities, and behavior management approaches that focus on gradually shaping the child's behavior to become more appropriate and productive in the classroom. Again, the teachers' beliefs about the degree of control that the child has can be an important first step in the strategies they develop for classroom organization and management.

## Methods

### Subjects

Information on 224 ED students was collected from 40 classroom teachers. The students came from 14 elementary schools, 14 middle schools, and 12 high schools in the city. There were 179 boys and 45 girls in the sample. The average age of the students was 12.1 years with a range of 7 to 19. There were 18 American Indian, 10 Black, 59 Hispanic, 130 White, and 7 "other" students. The students live in a variety of home settings -- 116 are in two-parent homes (74 with the biological parents and 42 in step-families); 82 children live with one parent (76 with mother and 6 with father); 26 of the children live in some other type of home situation. The average student was characterized as living in a lower middle class to middle class home.

Of the 40 teachers participating in the study, 28 were female and 12 were male. Thirty-seven of the teachers were White, 2 were Black, and 1 was Hispanic. The average age of the teachers was 30.2 with a range of 22 to 58. Twenty-seven of the teachers were married, and 13 were single at the time of the study. Approximately 40% of the teachers had Bachelor's degrees or were working on a Master's degree. The other 60% had Master's degrees or higher. The teachers averaged 2.8 years of experience with ED children with a range of 1 to 13 years.

## Research Methods

All of the information for the study came from four teacher questionnaires and rating scales. No testing or psychological information was available from school records. Anonymity was assured by the use of code numbers for both teachers and the students. Following is a description of each of the instruments used as well as the methods by which the information was compiled.

### Teacher Information Form.

This form asked for basic demographic information including sex, age, ethnicity, training, marital status, and years of experience.

### Pullis Inventory of Teacher Stress (PITS).

This is a three-part questionnaire that addresses areas concerning (a) sources of stress, (b) effects of stress, and (c) methods of coping with stress. Each section has several self-rating items as well as a question that invites the teacher to make additional comments. The items from the sources section ask teachers to rate four types of factors that have been found to contribute to teacher stress -- pupil characteristics, amount of work, career factors, and in-school conditions. The teacher is asked to indicate how stressful each factor is.

The section on the effects of stress is comprised of items describing typical thoughts, feelings and physiological problems that often accompany stress. For each item, the teacher is asked how often they experience the reaction described. Each teacher was given an overall score for this section that was based on summing all of their ratings across the items. This score was used in several of the analyses. It was felt that if teachers marked more items as occurring at a higher frequency, then they were likely to be experiencing more stress. It should be noted that this procedure is not a clinically verified measure of teacher stress. Rather, it is only a rough measure of teachers´ perceptions of these symptoms. The possible range of scores was 18 to 108. The average score for the sample was 44.05 (S.D. = 16.03) with a range of 24 to 90. These figures show the wide variability across the teachers. The average score means that most teachers marked most items as occuring "once in a while".

The final section asked teachers to indicate what strategies they use to cope with occupational stress as well as to rate each strategy according to its efficacy. Teachers were also asked to describe strategies that their school or district could adopt that they felt could reduce stress or assist them in dealing with stress.

Student Profile. This questionnaire had three parts that the teachers completed for each of the students in their class. The first part was the Quay-Peterson Behavior Problem Checklist (BPC). This is a 55-item list of behavioral problems commonly evidenced by ED students. Teachers are asked to indicate if the child has exhibited this problem. This scale has been used to determine four dimensions or types of behavioral difficulties -- conduct problems, withdrawn (neurotic), immaturity/inadequacy, and delinquency. These four types are determined by scores on particular items. While students sometimes receive high scores across different types, in most cases they tend to fall clearly into one of the four groups. For purposes of data analyses we used scores on each child from each of the dimensions or types. Thus, higher scores on any dimension mean that the child has evidenced more problems of that particular nature.

The second part of the Student Profile consisted of four items concerning the student´s intelligence, general level of motivation, social skills, and academic performance. These are traditional measures of student competence that are used in many studies. In past studies, we have found that teacher ratings of intelligence and academic performance correspond very accurately with actual IQ test scores and achievement test scores (Pullis & Cadwell, 1982). Since no testing information was available, it was felt that these ratings could be used with some degree of confidence when trying to examine the child´s ability and school performance and how those factors might be related to the child´s behavioral problems. In addition, when examining the section on teachers´ decisions, one would want to know generally how the child´s intelligence and academic progress influence teachers´ strategies. In this way, it can be determined how significant the learning problems are as compared to the behavior problems.

The third part of the student profile consisted of a 23-item temperament questionnaire. These items cluster into three factors or dimensions. The first is task orientation which refers to the student's ability to maintain a reasonable level of activity during tasks, persist on tasks until completion, and not be overly distracted. Adaptability refers to a child's tendency to approach new situations or people (as compared to a withdrawing tendency), to make reasonable adaptations to routine changes in the environment, and to evidence a positive quality of affect or mood during social interactions. Most of the items in this factor are set in a social context within school as compared to the task orientation factor. The third factor is called reactivity. This is essentially a factor that characterizes a student's intensity of behavior, their threshold of responsiveness, and negative aspects of mood during interactions. The items are weighed in such a way that higher scores on task orientation and adaptability are considered to reflect more positive performance. Lower scores on the reactivity dimension mean the student is overly intense, easily over-stimulated and conveys a negative quality of mood. Thus, from the Student Profile there are 11 measures of students' attributes or characteristics which were used throughout the data analyses.

### • Dimensions of Behavior Disorders

| | |
|---|---|
| Conduct Problem Ratings | The higher the scores on each of these dimensions, the more pervasive the problems (higher frequency of problem checked by the teacher). |
| Withdrawn Ratings | |
| Inadequate/Immaturity Ratings | |
| Delinquency Ratings | |

### • Student Competence

| | |
|---|---|
| General Ability or Intelligence | The higher the score on each of these variables, the more positive the performance. |
| Classroom Motivation | |
| Social Interaction Skills | |
| Academic Achievement Performance | |

### • Temperament Ratings

| | |
|---|---|
| Task Orientation | Higher ratings mean more positive performance, or more positive behavioral attributes. |
| Adaptability | |
| Reactivity | |

Classroom situations and decisions. This instrument was designed to examine various aspects of classroom interactions to determine how teachers respond to students and to document different decisions that they make regarding classroom management. There were 14 different decisions that were tapped in this section. Seven of the decision concerned how often teachers had to monitor students across different classroom situations. These situations were: (a) independent seatwork, (b) teacher presentation or lecture to the entire class, (c) small group activity directed by students, (d) teacher-directed, small group lesson, (e) academic transition -- changing from one in-class activity to another, (f) nonacademic transition -- changing from an activity outside the classroom to one inside the classroom, and (g) free-time activity. These situations were chosen because they are typical to most classrooms. The monitoring decision was asked as a way of measuring the teacher's anticipatory decision that a child would require close monitoring during the various situations. Research has shown that those teachers who anticipate or who are proactive tend to manage problems more efficiently through restructuring or defusing escalating problems. Thus, it was felt that this was an important decision to measure. On these questions, a higher rating means the teacher feels that the child needs monitoring more often during the classroom activity.

The next three decisions were considered instructional management strategies. Teachers were asked to indicate how often they have to (a) change the child's seat location during instructional times, (b) move to the child or move the child to them to provide extra help on a task, and (c) modify the content or methods by which they teach each particular child. These decisions reflect those differential strategies teachers often have to use with ED students. Higher ratings on these items mean that the teacher has to use the strategies more often in order to effectively manage the child during instructional times.

The next three items were designed to examine the teachers' approaches to behavior management. The teachers were first asked to indicate (from an ordered list) the type of reinforcement they found to be most effective with each student. The choices included edibles, tangible rewards, awards, tokens, free-choice time, and teacher praise. Next, the teachers were asked to indicate the strategies they used when a child was inappropriate or disruptive in the classroom. They were asked to mark (again from an ordered list) what they typically used and then were asked to note the most directive management strategy they used with a particular child. The list ranged from ignoring and nonverbal techniques to removal and corporal punishment. The higher the rating here, the more directive the management techniques that the teacher felt they had to use.

The final item requested that the teachers make a placement recommendation for each student for the coming school year. There were five choices that ranged from full-time special education to full-time mainstreaming into a regular class.

Control. In addition, teachers were asked to rate how much control they felt each child had concerning their behavior. They were prompted to consider such things as strong family influences, biological or neurological factors, etc. For the purposes of this study this rating was used to determine how much responsibility (control and effort) the teacher felt each child was assuming concerning their school behavior and academic performance.

## Results and Discussion

Results from two analytic procedures will be presented here. First, correlational analyses were conducted to determine if there were significant relationships between the teachers' perceptions of control and other ratings that they had completed on their emotionally disturbed students. The results of this analysis are presented in Table 1. Second, a multiple regression or policy capturing analysis was conducted to determine those factors that are predictive of teachers' classroom decisions and strategies. These results are presented in Table 2.

## Table 1

### Correlations Between Teacher Ratings of Control and Demographics and Competence Ratings, Classroom Decisions, and Temperament and BPC Ratings

| Student Information | Correlation |
|---|---|
| **Demographics, Competence Ratings, and Control** | |
| Age | .02 |
| SES | .01 |
| Home Influence | .27*** |
| Intelligence Estimate | .21** |
| Motivation | .19** |
| Social Skills | .49*** |
| Academic Performance | .18** |
| **Classroom Decisions and Control** | |
| Independent seatwork | – .25*** |
| Whole class lesson | – .30*** |
| Group activity | – .31*** |
| Small group lesson | – .29*** |
| Academic transition | – .31*** |
| Nonacademic transition | – .35*** |
| Free time | – .22** |
| Change seat location | – .29*** |
| Move to assist child | – .29*** |
| Modify instruction | – .32*** |
| Reinforcement strategy | .06 |
| Management strategy | .35*** |
| Directive management | .48*** |
| Placement decision | .35*** |
| **Temperament, BPC Ratings, and Control** | |
| Task orientation | .32*** |
| Adaptability | .20** |
| Reactivity | .19** |
| Conduct problems | – .27*** |
| Personality disorders | – .07 |
| Immaturity | – .16* |
| Delinquency | – .02 |
| **Teacher Stress and Control** | .44*** |

Note. *** $p < .0001$
** $p < .001$
* $p < .01$

### Table 2
### Multiple Regression Findings Showing Factors That Predict Significantly Teachers' Decisions

| Decision Situations | Significant Factors[a] |
|---|---|
| Monitoring Decisions | |
| Independent seatwork | – Task orientation, – Academic performance, –Motivation, – Reactivity |
| Whole class lesson | – Task orientation, – Motivation |
| Group activity | – Task orientation, –Conduct problems, –Social skills |
| Small group lesson | – Task orientation, –Motivation |
| Academic transition | – Task orientation, – Adaptability, – Immaturity, – Reactivity, – Control |
| Nonacademic transition | – Task orientation, – Reactivity, – Motivation, – Delinquent, – Control |
| Free Time | – Task orientation, – Delinquent, – Conduct problems, + Adaptability |
| Strategies | |
| Change seat location | – Task orientation, – Social skills, – Conduct problems |
| Move to assist child | – Task orientation, – Intelligence estimate |
| Modify instruction | – Intelligence estimate, – Task orientation, – Personality disorders |
| Reinforcement strategy | + Conduct problems, + Task orientation |
| Management strategy | No significant predictor variables |
| Directive management | No significant predictor variables |
| Placement decision | + Conduct problems, + Task orientation, + Social skills, – Reactivity, + Control |

---

[a] The factors are listed in the order of magnitude based on regression weights. Only those factors that were statistically significant are listed. Signs preceeding the factors indicate direction of relationship. For example, the negative sign preceeding "Control" in the Academic Transition monitoring decision means that as the child was rated lower on the degree of control they have over their behavior, the higher the need to monitor the child.

The correlational analysis revealed several interesting and important findings concerning the teachers' perceptions of control and other aspects of student characteristics as well as teachers' decision strategies. It should be noted here that because of the large sample size, small to moderate correlations can achieve rather high levels of statistical significance. Thus, it is probably most appropriate to think of these findings as indicative of trends that may be investigated in the future with more rigorous designs and statistical procedures.

## Demographics and Competence Ratings

Rather surprisingly, there was no significant relationship between the ratings of control and the factors of age and SES. It had been anticipated that older students would be rated as having more control over their behavior and that poorer students would be rated as having less control because of disadvantaged lives, cultural factors, etc. This was not borne out in the correlational findings. There was a positive relationship between the teachers' ratings of home influence and their ratings of control. The more positive the home influence, the greater degree of control the students were perceived to have over their behavior. The quality of home influence was not related to SES level.

The ratings of degree of control were positively and significantly correlated to the ratings of intelligence, motivation, social skills, and academic performance. Thus, children rated on all of these general competence ratings were perceived to have a greater degree of control over their behavior.

## Classroom Decisions

There were two significant trends that emerged from the correlation analysis concerning teachers´ classroom decisions. The first trend was that ratings of control were significantly negatively correlated with the monitoring and instructional decisions of the teachers. This means that lower ratings concerning degree of control were associated with higher ratings concerning the frequency of monitoring during a variety of classroom activities, changing the child´s seating, helping the child, and modifying instructional activities. It appears that children who were perceived as having less control over their behavior required additional teacher time and energy during classroom interactions. This finding is quite logical as children with less control would need more direction and assistance.

The other trend concerns teachers´ responses to children´s behavior. There was no relationship between degree of control and the typical reinforcement strategy utilized by the teachers. However, there were significant positive correlations between the ratings of degree of control and the discipline strategies used by the teachers. Recall that teachers were asked to indicate on a hierarchy of 10 management approaches both the typical strategy they used and the most directive strategy they used when the child was disruptive or not on-task in the classroom. The higher the ratings on this hierarchy, the more directive or intrusive the management or discipline approach. The positive correlation here means that children perceived as having more control were responded to with more directive discipline approaches. This finding appears to confirm Kauffman´s contention that these children may elicit more negative feelings from teachers that are then translated into more harsh or directive management strategies. Caution in the interpretation here is warranted, but it may be that this find reflects the negativity that can develop when the teachers believe that the child does have control over their behavior but is not putting forth sufficient effort. Finally, higher ratings of control were positively correlated with higher placement recommendations. It should be noted here that almost all of the teachers´ recommendations were for continued enrollment in self-contained ED programs. However, some children were thought to be ready for less restrictive resource programs. The positive correlation here means that children viewed as having more control might be ready for less restrictive placements in the coming school year.

## Temperament and BPC Ratings

There were positive correlations between the ratings of control and all three aspects of temperament -- task orientation, adaptability, and reactivity. Children perceived as having more control over their behavior also received higher or more positive temperament ratings.

Ratings on the Behavior Problem Checklist yielded scores on four aspects of disordered behavior -- conduct problems, personality disorders, immaturity, and delinquency. Higher ratings on this measure mean that more items were checked by the teacher. There were no significant correlations between perceptions of control and ratings of personality disorders and delinquency. There were significant negative correlations between ratings of control and the number of items checked in the areas of conduct disorders and immaturity. These correlations mean that the presence of more problems in these two areas was associated with ratings of less control on the part of students. In the case of conduct disorders, that may mean that children evidencing a number of problematic behaviors are viewed as having less control over their behavior. It may be that teachers believe these children come from troubled homes or that there may be some biological or neurological factors influencing these children´s behavior. This same trend or belief may also be true in the area of immaturity where the teachers sense that the child´s behavior may reflect some cognitive or developmental delay.

The final finding of note regards the positive correlation between ratings of control and the teachers´ reports concerning symptoms of stress. In the section on stress, teachers were asked to note the frequency with which they experienced certain feelings (either psychological or physiological) that past research has associated with anxiety reactions to stressful work situations. It is interesting that higher reported frequencies of stress reactions were positively correlated with higher ratings of degree of control. This finding may indicate that working with problematic students who are perceived to have control over their behavior may be more stressful for the teachers. Taken together with the findings on the use of more directive management approaches with these children, this correlation suggesting higher stress levels may be reflective of the negative interactions between teachers and children who teachers feel could be behaving more appropriately.

The other analysis used in this study was a multiple regression or policy capturing procedure. In this technique, the goal was to determine which factors were predictors of the teachers´ decisions. The role of temperament in these decision-making situations has been discussed elsewhere (Pullis, 1983). This procedure was also used to determine if the ratings of control were significantly related to the teachers´ decision when considered simultaneously with other factors like student characteristics, demographic information, etc. While there were consistent trends in the correlational analysis associating ratings of control with teacher decisions, this analytic procedure could be considered a more stringent test as teachers do consider several factors in formulating their decisions.

Looking over the results presented in Table 2, one can see that the ratings of control were statistically significant in only three decisions. Lower ratings of control were predictive of more frequent needs to monitor students in transition situations. This is a common problem reported by teachers of emotionally disturbed students. Structure tends to break down during times when children are switching learning activities (academic transitions) or moving in and out of the room (nonacademic transitions). Higher ratings of control were predictive of higher (less restrictive) placement recommendations for the coming school year.

Findings from the correlational and multiple regression analyses do lend some support to the notion that teachers´ perceptions concerning the amount of control each student has over their behavior problems may be an important mediating factor that influences the teacher-student relationship as well as specific strategies or decisions that the teachers make. As suggested earlier, this study has focused only on one specific aspect of teacher thinking and it should be remembered that several other teacher ratings were garnered during the total research project. These findings should be considered as preliminary evidence about teacher cognitions concerning their students´ behavior. Other methods should now be employed to investigate more systematically the important role of teachers´ beliefs and interpretations concerning their pupils´ behavioral problems and causes of those problems. The emerging literature on teacher cognitions would be a useful place to begin examining experimental and observational methods for this more rigorous approach to research in this area.

## Acknowledgment

The gracious cooperation of the Albuquerque, New Mexico Public Schools is acknowledged.

## References

Brophy, J., & Rohrkemper, M. (1981). Influences of problem ownership on teachers´ perceptions of and strategies for coping with problem students. *Journal of Educational Psychology*, *73*, 295-311.

Kauffman, J. (1981). *Characteristics of children´s behavior disorders*. Columbus: Merrill.

Keogh, B., Pullis, M., & Cadwell, J. (1982). A short form of the teacher temperament questionnaire. *Journal of Educational Measurement*, *19*, 323-329.

Pullis, M. (1983, April). *Temperament and behavior disorders: Influences on teachers´ perceptions and classroom decisions*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Pullis, M. (in press). LD students´ temperament characteristics and their impact on decisions made by resource and mainstream teachers. *Learning Disability Quarterly*.

Pullis, M., & Cadwell, J. (1982). The influence of children´s temperament characteristics on teachers´ decision strategies. *American Educational Research Journal*, *19*, 165-181.

Pullis, M., & Cadwell, J. (in press). Temperament as a factor in the assessment of children educationally as risk. *Journal of Special Education*.

Thomas, A., & Chess, S. (1977). *Temperament and development*. New York: Brunner/Mazel.

Thomas, A., Chess, S., & Birch, H. (1968). *Temperament and behavior disorders in children*. New York: New York University Press.

# AN ANALYSIS OF INDIVIDUAL DIFFERENCES IN STUDENT TEMPERAMENT CHARACTERISTICS AND THE IMPLICATIONS FOR CLASSROOM PROCESSES AND OUTCOMES

Lizanne DeStefano and Margaret C. Wang
University of Pittsburgh

and

Edmund W. Gordon
Yale University

The purpose of this paper is to report the findings from a descriptive study of the instructional implications of learner differences in temperament characteristics. The paper is organized in four major sections. The first is a brief discussion of the conceptual framework and rationale for the work described here. The study's design and the results on the effects of temperament characteristics on students' classroom processes and learning outcomes are presented in the next two sections. The final part of the paper provides a discussion of the implications of the study's findings for improving educational programming through use of information on differences in temperament characteristics to maximize learning outcomes for individual students.

## The Conceptual Framework

The study described in this paper was based on the contentions that temperament is one of the particular clusters of indigenous and acquired characteristics which are unique to individual learners, and that the interaction of these learner characteristics with the learning environment is significantly related to the learning behaviors and outcomes of individual students. Discussion in this section focuses on the nature of the relationship between temperament and learning as it has been hypothesized and reported in the literature, as well as the specific model of temperament characteristics and school learning that provided the framework for the present study.

## Temperament and Learning

The premise that temperament has a direct and indirect influence on the interactions of individual students with their learning environment is based on a view of temperament characteristics as an individual difference variable. Buss and Plomin (1975) suggest that temperament directly influences learner behavior by helping to determine the ways in which individuals initiate interactions, the ways in which they structure and interact with their environment, and the ways in which others perceive and respond to them. Work by Kagan and Kogan (1970) also points to the direct effect of individual difference variables such as cognitive style and temperament characteristics on behavior, which, in turn, influences performance.

The indirect influence of temperament on student learning is suggested from research using the expectation model of Brophy, Good, and Evertson (Brophy & Evertson, 1981; Brophy & Good, 1974; Good & Brophy, 1973). This model is based on the contention that teachers' perceptions of, and reactions to, patterns of temperament may result in certain teacher expectations for particular students. Pullis (1979) supports this line of thinking with his finding that teachers regularly overestimated the ability of students with positive temperament patterns while often underestimating that of students with more negative temperament attributes.

In addition to the effects on teacher perceptions and subsequent teacher-student interactions, variations in temperament patterns influence learner behaviors such as the amount of time spent on-task, the degree of task involvement, and the use of classroom resources and options. In turn, research findings have suggested a relationship between these learner behaviors and teacher perceptions and the extent and nature of learning outcomes such as achievement and adjustment (Berliner & Rosenshine, 1976; Bloom, 1968; Denham & Lieberman, 1980; Mischel, Zeiss, & Zeiss, 1974; Rutter, Maughan, Mortimore, & Ouston, 1979; Wang & Stiles, 1976). Yet few studies on the instructional implications of individual differences have examined either the effects of temperament characteristics on students′ behavior and adjustment in school situations or the nature of the influence of temperament on students′ achievement, motivation, and perceptions of their own self-competence.

Although research does not suggest that temperament characteristics are the direct or single cause of learning and behavior problems, there is empirical evidence to support the hypothesis that a child′s temperamental predisposition influences his or her personal-social transactions by reducing or intensifying potential problems (Graham, Rutter, & George, 1973; Rutter, Birch, Thomas & Chess, 1964; Terestman, 1980). A number of studies have examined temperament-school compatibility. For example, in the New York Longitudinal Study, Thomas, Chess, and Birch and their associates (1963, 1968) found that Difficult child temperament patterns were highly associated with behavior problems, and Slow-To-Warm-Up temperament patterns were related to school achievement problems. Other research (Carey, Fox, & McDevitt, 1977;

Hall & Keogh, 1978; Lambert & Windmiller, 1977) has associated similar patterns of temperament characteristics with patterns of academic achievement and adjustment.

Thomas, Chess, and Birch (1968) use the term "goodness of fit" to suggest that the course of a child's development is, in part, a function of the extent of consonance or dissonance between the child's temperament and its parents' child-rearing practices. However, goodness of fit is not limited to the family situation. There are varying degrees of compatibility between the temperamental predisposition of an individual student and the structure of school programs or the nature of personal-social demands in the classroom. This degree of compatibility is manifested in the student's learning behavior, achievement, adjustment, and competence. The study described in this paper was designed to explore the applicability of the goodness of fit notion in school learning situations by investigating the interactive effects of temperament characteristics, school learning environments, student behavior, and student learning outcomes.

## The Model of Temperament Characteristics and School Learning

Figure 1 is a schematic representation of the Model of Temperament Characteristics and School Learning (Gordon & Wang, 1982) that served as the framework for the present study. As shown in the figure, the Model consists of four major clusters of variables -- learner characteristics, learning conditions, learner behaviors, and learning outcomes -- and the hypothesized interactive relationship among them. The premise underlying the design of the

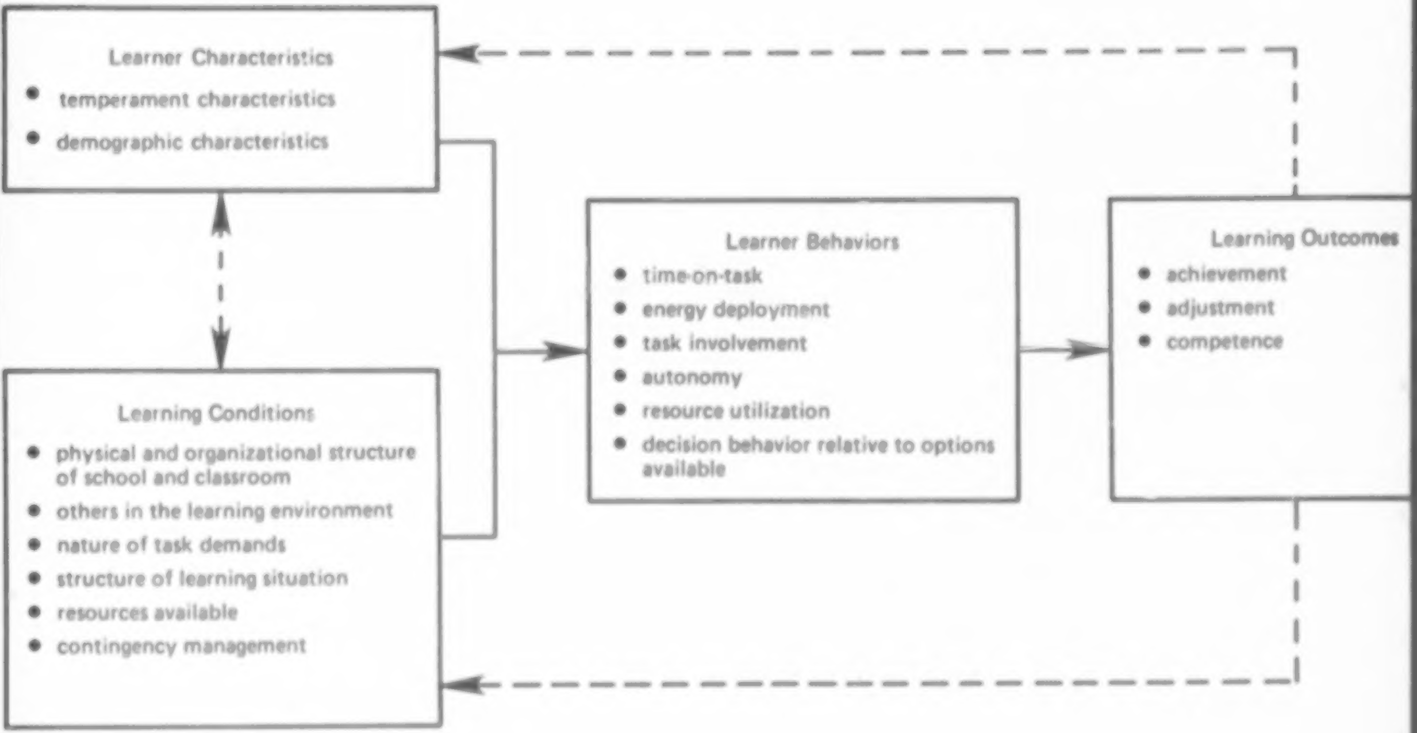
Learner Characteristics
• temperament characteristics
• demographic characteristics
Learning Conditions
• physical and organizational structure of school and classroom
• others in the learning environment
• nature of task demands
• structure of learning situation
• resources available
• contingency management
Learner Behaviors
• time-on-task
• energy deployment
• task involvement
• autonomy
• resource utilization
• decision behavior relative to options available
Learning Outcomes
• achievement
• adjustment
• competence

Figure 1. Model of Temperament Characteristics and School Learning.

BEST COPY AVAILABLE

Model is that the individual learner brings to the school learning environment an unique profile of instructionally relevant temperament characteristics. These characteristics interact with learning conditions, such as the physical and organizational structure of the classroom, the nature of task demands, and available instructional-learning resources, to elicit certain learner behaviors. The learner behaviors included in the Model are time-on-task, energy deployment, task involvement, autonomy, resource utilization, and decision-making behavior. These learner behaviors are hypothesized to be related to learning outcomes that include academic achievement, and these outcomes, in turn, are expected to interactively affect the temperament characteristics of individual students.

## The Study

The present study was designed to investigate the hypothesized relationship among the four clusters of variables shown in Figure 1. The following series of questions was addressed.

1. What identifiable patterns of temperament characteristics are present in the study's sample of primary-grades students?

2. What differences exist among the identified patterns of temperament characteristics and student demographic characteristics such as age, sex, ethnicity, SES, and prior achievement?

3. What differences exist among the identified patterns of temperament characteristics and student learning behaviors such as time-on-

task, the nature and patterns of student-teacher interactions, the types of learning tasks on which students tend to work, and the manner in which classroom time is spent?

4. What differences exist among identified patterns of temperament characteristics and learner achievement, as measured by task completion rates, number of curriculum objectives mastered, and scores on standardized achievement tests in reading and math?

5. What differences exist among identified patterns of temperament characteristics and students′ and teachers′ perceptions of the adjustment and competence outcomes for individual learners?

### Setting

The study was carried out during the 1982-83 school year in a multi-aged, primary-grades classroom of a university laboratory school. The population of the classroom consisted of 66 students ranging in age from 4 to 9, two teachers, one student intern, and one instructional aide.

An instructional program known as the Adaptive Learning Environments Model (ALEM) was implemented in the classroom. The central goal of the ALEM is to provide learning environments that are responsive to the characteristics and learning needs of individual students (Wang, 1980a). The program′s design systematically integrates aspects of individualized prescriptive instruction that facilitate basic skills mastery (Bloom, 1976; Glaser, 1977; Rosenshine, (Johnson, Maruyama, Johnson, Nelson, & Skon, 1981; Marshall, 1981; Peterson, 1979). The ALEM consists of 12 critical program dimensions. Nine are related to the process of providing adaptive instruction. They are Creating and Maintaining Instructional Materials, Developing Student Self-Responsibility, Diagnostic Testing, Instructing, Interactive Teaching, Monitoring and Diagnosing, Motivating, Prescribing, and Record Keeping. Three of the dimensions -- Arranging Space and Facilities, Establishing and Communicating Rules and Procedures, and Managing Aides -- are related to providing support for implementation of adaptive instruction in the classroom. In addition to the 12 dimensions, implementation of the ALEM is supported by a delivery system that is made up of four major components: an ongoing, data-based staff development approach; instructional teaming; multi-age grouping; and family involvement. (For a more detailed description of the ALEM, see Wang, Gennari, and Waxman, 1985.)

The ALEM classroom was chosen as the site for the present study for three reasons. First, the program's aim to accommodate diverse student learning needs is more likely to permit individual differences to surface than would conventional group-instruction programs. Thus, a good opportunity existed to study the relationships between individual-difference variables. The second reason for choosing to conduct the study in this particular classroom was that it represented a mainstreaming setting where exceptional and general education students were integrated. The result was a heterogeneous population that included mildly handicapped students and gifted students with special learning needs. The final reason for selection of the study site was the availability of a great deal of individual-difference data through the ongoing assessment of program implementation and student outcomes in the classroom.

## Measures and Procedures

A variety of techniques and instruments was used to obtain data for the study. They included observations in classroom settings, questionnaires and rating scales, structured interviews, and school records. These are described briefly below for each of the four clusters of variables included in the study (learner characteristics, learning conditions, learner behaviors, and learning outcomes).

In order to record entry and exit characteristics, as well as capture the amount and nature of change in the major categories of variables over time, each type of data was collected two or more times during the school year (fall, winter, spring). Only students with completed measures for all the collection periods were included in the analysis (N = 56).

Learner characteristics. Data on this variable consisted of information on students' demographic and temperament characteristics. Demographic data were collected for all students in the sample population at the beginning of the school year. This information was updated at mid-year and again at the end of the school year. School records and teacher reports provided information on age, sex, ethnicity, SES, and dominant language. They also served as a source of information on students' prior achievement. The latter data included scores from the previous school year on standardized achievement tests in reading and math, progress in the ALEM's reading and math curricula, and average task completion rates in reading and math.

Keogh's (Keogh, Pullis, & Cadwell, 1982) short form of the Teacher Temperament Questionnaire (TTQ) developed by Thomas and Chess (1977) was used to measure students' temperament characteristics. The design of the short form is based on a series of factor analyses of the original 64-item TTQ. Although there is a great deal of research evidence supporting the validity of the original TTQ (e.g., Billman & McDevitt, 1980; Carey, 1981; Maurer, Cadoret, & Cain, 1980), its length limits its utility in large-group educational settings. Thus, Keogh and her associates focused on using the TTQ to create a measure of temperament that is both psychometrically adequate and practically feasible for use by teachers. The short form is a 23-item scale representing all eight dimensions of the original instrument. The results of Pullis' (1979) psychometric analyses of the short form suggest its utility and reliability in assessing teachers' perceptions of students' temperament characteristics.

The short form of the TTQ requires approximately 15 minutes per student to administer and a minimum amount of training for teachers. In the present study, training consisted of explaining the categories or dimensions of temperament and their behavioral manifestations and giving a brief overview of the implications of temperament characteristics for student learning. The short form was administered twice, in fall and spring of the school year. Each student in the sample was rated by his or her homeroom teacher. A random sample of 50% of the students was rated by both teachers in the class (each of whom was familiar with the students), and a correlation coefficient was computed to determine the strength of agreement between teachers´ ratings for each temperament dimension and the overall scores. Inter-rater reliability was high at .63. The fall-spring, rate-rerate reliability was .57.

As a way of examining the extent to which the three-factor structure formulated by Thomas and Chess (1977) and Keogh (1982) was present in the sample of information collected in the present study using the short form of the TTQ, individual item scores and subscale scores across the sample were factor-analyzed. Overall, the same three factors were verified, and together they accounted for 63% of the variance in the sample.

The first factor, task orientation, was loaded mainly on items from the three dimensions, activity level, persistence, and distractibility. The second factor, personal-social flexibility, was loaded mainly on items from the dimensions, adaptability, approach-withdrawal, and (positive) quality of mood. Finally, the factor, reactivity, was constituted of items from the dimensions, intensity of reaction, threshold of responsiveness, and (negative) quality of mood. Teachers´ ratings for individual students on all items of the TTQ were combined to give average scores for each of the three temperament factors. For example, students´ ratings on items 1,2,3,9,10,11,17,18, and 19 were added together to give a mean task orientation score for each student.

Based on each student´s profile on the three factors, the students were grouped into the three temperament categories developed by Thomas and Chess (1977). These categories were Easy, Difficult, and Slow-To-Warm-Up. Easy students were characterized by high task orientation, high personal-social flexibility, and low-to-average reactivity; Difficult students by low task orientation, low-to-moderate personal-social flexibility, and high reactivity; and Slow-To-Warm-Up students by low-to-moderate task orientation, low personal-social flexibility, and moderate-to-high reactivity. Rankings such as low, moderate, and high were determined by calculating the deviation of each student's factor score from the mean factor scores for the individual teachers.

Learning conditions. The Implementation Assessment Battery for Adaptive Instruction (Wang, 1980b; Wang, Catalano, & Gromoll, 1983) was used to describe the learning conditions under which educational activities were implemented in the classroom. A total of 106 performance indicators, grouped into the Battery of six data collection forms, have been identified to assess the presence and absence of the 12 critical program design dimensions of the ALEM. Results from an empirical validation study of the Battery suggest its reliability and validity (Strom & Wang, 1982): The overall generalizability coefficient was .88, and the generalizability coefficients for the 12 dimensions ranged from .50 (Record Keeping) to .91 (Interactive Teaching).

Administration of the entire Battery requires approximately two hours. It was used by trained observers in the present study to collect degree of implementation data at three different points during the 1982-83 school year.

Learner behaviors. The Student Behavior Observation Schedule (Wang, 1974b) was used to obtain information on classroom process patterns. The Student Behavior Observation Schedule (SBOS) has been used in a number of investigations of classroom processes under the ALEM, with inter-observer agreement consistently found to be above 85% (Wang, 1976). Of particular interest in the present study were students' time-on-task, the nature and patterns of interactions between teachers and students, the settings in which learning activities occurred, the types of learning tasks on which students chose to work, and the manner in which classroom time was spent.

All students in the study were observed by trained observers using the SBOS. Each observation consisted of 10 alternating, one-minute intervals per student, distributed over a 50-minute time block. The SBOS was administered at three different points (fall, winter, spring) during the school year.

Learning outcomes. Data were collected on three categories of learning outcomes -- achievement, adjustment, and competence. School records of progress in basic academic skill areas were the primary source of data on learner achievement. Information on learning progress included (a) scores on math and reading subtests of standardized achievement tests administered in fall and spring of the school year; (b) the number of curriculum objectives mastered in math and reading by the end of the school year; and (c) task completion rates (at three points in the school year) based on classroom records of the numbers of tasks assigned by teachers and the numbers of tasks completed by individual students.

The learning outcomes, adjustment and competence, were assessed through information on student and teacher perceptions. The Self-Responsibility Interview Schedule (Wang, 1974a) and the Perceived Competence Scale (Harter,

1982) for children were used to collect data on student perceptions. Information on teacher perceptions was obtained through the Teacher Perception Questionnaire (Weisstein & Wang, 1980).

The Self-Responsibility Interview Schedule (SRIS) is designed to assess students' knowledge of the operational aspects of the learning environment as well as their perceptions of their own competence and adjustment for functioning in the environment. The SRIS consists of 21 questions aimed at obtaining three categories of information: students' knowledge about the learning environment, their ability to evaluate their own learning, and their perceptions of personal control over their own learning. Particular attention was given in the present study to students' responses to questions such as, "Do you think that you will finish your work today?"; "Do you prefer doing reading and math with everyone else when the teacher tells you to, or deciding for yourself when to do reading or math or other work?"; and, "Do you think that you are a hard worker?" The SRIS was administered to students individually by trained interviewers who read the prepared list of questions and recorded the students' responses. Each interview took approximately 10 minutes to administer and score. SRIS data were collected on all students in the first through third grades during fall and spring of the school year.

The Perceived Competence Scale (PCS) for children was used to assess students' perceptions of their own cognitive, social, and physical competence and their overall feelings of self-esteem. The PCS uses a structured rating-scale format. Students are asked to describe how similar they think they are to hypothetical children. Their responses are rated on a scale from

1 to 4, with 4 being the most positive response (e.g., feeling pretty sure of oneself). The PCS takes approximately 10 minutes per administration. It was administered to all first-, second-, and third-grade students in the study sample in fall and spring.

Information on teachers' perceptions of students' adjustment and competence was obtained through the Teacher Perception Questionnaire (TPQ), which is designed to elicit teachers' judgments of individual students as high, medium, or low along a number of academic and social dimensions. Specifically, data for the present study included information on teachers' perceptions of each student in terms of (a) level of academic achievement, (b) effort expended on instructional tasks, (c) ability to work independently, (d) number of instructional tasks the student was capable of completing, and (e) popularity in the learning environment. Teachers' perceptions of the amount of control they exerted over students in their class also were measured. The TPQ was administered to teachers individually. For each of six questions, a set of index cards was spread in front of a teacher, each with the name of one student in his or her classroom. The teacher was asked to sort the cards into three piles according to his or her perception of individual students as high, average, or low in a particular area compared to other students in the same classroom. The TPQ was administered once in the fall and once in the spring of the school year.

## Results of the Study

The results of the study are summarized in this section according to the specific questions that the study was designed to answer. Findings are reported on the identifiable patterns of temperament characteristics for students in the classroom and the differences among students with different patterns of temperament characteristics in terms of demographic characteristics, learner behaviors, learner achievement, and adjustment and competence.

It should be noted here that the mean degree of implementation scores for the 12 critical dimensions of the ALEM, across all three data collection periods, were above the 85% criterion for a high degree of implementation. Thus, the learning conditions under which the students worked represented a highly adaptive school learning environment as specified by the design of the ALEM.

## Patterns of Temperament Characteristics

Based on factor analyses of results from the Teacher Temperament Questionnaire (TTQ), 23 (38%) of the students in the study sample were categorized as Easy; 13 (22%) were categorized as Difficult; and 20 (33%) were categorized as Slow-To-Warm-Up. Four students (7%) were Unclassified; that is, they did not fit into any of the three temperament categories.

The students' temperament ratings on the eight dimensions measured by the TTQ are summarized in Table 1 according to the three temperament categories. Statistically significant differences were noted for five of the eight dimensions. Non-significant differences among the three temperament categories were found for the dimensions, approach-withdrawal, threshold of responsiveness, and distractibility. The differences among the groups along

### Table 1

### Mean Scores From the Teacher Temperament Questionnaire (TTQ) for Students in the Three Temperament Categories

| | Dimensions of the TTQ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Temperament Category | Activity Level | Adaptability | Approach-Withdrawal | Threshold of Responsiveness | Intensity of Reaction | Quality of Mood | Distractibility | Attention Span and Persistence |
| Easy<br>(N = 23) | 8.34<br>(3.30) | 15.37<br>(3.09) | 10.17<br>(4.73) | 6.60<br>(1.75) | 8.65<br>(4.57) | 18.43<br>(8.16) | 8.35<br>(5.75) | 20.00<br>(6.89) |
| Difficult<br>(N = 13) | 12.20<br>(3.19) | 10.60<br>(3.38) | 8.70<br>(3.70) | 7.00<br>(.96) | 11.77<br>(3.92) | 10.77<br>(2.28) | 10.54<br>(2.96) | 7.92<br>(4.31) |
| Slow-To-Warm-Up<br>(N = 20) | 7.50<br>(3.83) | 9.70<br>(2.70) | 8.35<br>(2.13) | 6.80<br>(1.85) | 7.90<br>(1.86) | 11.65<br>(4.48) | 9.85<br>(2.54) | 10.95<br>(6.02) |
| Results from tests of significant differences among the three temperament groups | $F=7.78$<br>$p \leq .01$ | $F=25.36$<br>$p \leq .01$ | $F=2.95$<br>$p \leq .10$ | $F=.51$<br>$p \geq .10$ | $F=9.42$<br>$p \leq .01$ | $F=36.43$<br>$p \leq .01$ | $F=2.57$<br>$p \leq .10$ | $F=37.71$<br>$p \leq .01$ |

each of the dimensions were found to be consistent with findings from the New York Longitudinal Study (Thomas, Chess, & Birch, 1968).

## Differences in Demographic Characteristics Across Temperament Patterns

Table 2 provides a summary of demographic information for students in the three temperament categories. As shown in the table, the distribution of students by sex showed a greater percentage of males than females in the Easy category (46% vs. 31%) and a much greater percentage of females than males in the Slow-To-Warm-Up category (41% vs. 25%). These findings are at variance with previous research which reports that girls are rated as having Easy temperaments about twice as often as are boys (Keogh, 1982; Thomas & Chess, 1977). A possible explanation for the differences in these findings may lie in the interaction between learner characteristics and learning conditions. Previous research was conducted in traditional classrooms where the degree of structure was greater and the freedom of movement more restricted. In the ALEM classroom where the present study was conducted, students were free to move about, select activities and work spaces, and interact with their peers. As a result, students who were highly active (a major dimension of the Difficult temperament category) had opportunities to channel that activity. In light of the research that has found boys to be more physically active than girls in the primary years (Dweck & Goetz, 1977; Meyer & Thompson, 1966), the discharge of activity in a socially appropriate manner may explain, at least in part, the higher than expected percentage of boys than girls in the Easy category.

Table 2

Summary of Demographic Characteristics
of Students in the Three Temperament Categories

| | Temperament Category | | | |
|---|---|---|---|---|
| Demographic Characteristics | Easy (N=23) 38% | Difficult (N=13) 22% | Slow-To-Warm-Up (N=20) 33% | Unclassified (N=4) 7% |
| **Sex** | | | | |
| Male | (13) 46 | (6) 21 | (7) 25 | (2) 8 |
| Female | (10) 31 | (7) 22 | (13) 41 | (2) 6 |
| **Grade Level** | | | | |
| Kindergarten | (0) 0 | (7) 37 | (12) 63 | (0) 0 |
| Grade 1 | (6) 35 | (4) 24 | (5) 29 | (2) 11 |
| Grade 2 | (8) 67 | (2) 17 | (1) 8 | (1) 8 |
| Grade 3 | (9) 75 | (0) 0 | (2) 17 | (1) 8 |
| **Ethnicity** | | | | |
| Caucasian | (16) 70 | (9) 71 | (14) 70 | (4) 100 |
| Black | (5) 22 | (2) 15 | (3) 15 | (0) 0 |
| Asian American | (1) 4 | (1) 7 | (0) 0 | (0) 0 |
| Other | (1) 4 | (1) 7 | (3) 15 | (0) 0 |
| **Dominant Language** | | | | |
| English | (21) 91 | (11) 85 | (17) 85 | (4) 100 |
| Other | (2) 9 | (2) 15 | (3) 15 | (0) 0 |
| IQ (mean score for each temperament category) | 117<br>(12.06)[a] | 112<br>(12.53) | 110<br>(13.08) | 114<br>(11.41) |

Note. [a] Numbers in parentheses under mean IQ scores represent standard deviations.

No significant differences were found among the temperament categories in terms of ethnicity, dominant language, and IQ. However, when the three temperament categories were broken down by grade level, a consistent trend was noted. As shown in Table 2, the proportion of Easy students increased with each higher grade level. Conversely, the percentage of Difficult students was greatest in Kindergarten, with lower percentages of students in this category in each higher grade.

It may be speculated that the greater percentages of Easy students in the higher grade levels were the result of cumulative program effects as the students learned to function more effectively in an adaptive learning environment. It has been hypothesized that in an instructional program designed to accommodate students' individual differences by allowing them to choose learning activities and settings and to plan their instructional day, the kinds of descriptors which characterize Difficult and Slow-To-Warm-Up students (low adaptability, limited approach-withdrawal, high reactivity) are likely to be modified and replaced by the more adaptive characteristics of the Easy student. Further research that addresses these intervention issues would be very useful in delineating program design requirements for fostering development of Easy temperament characteristics.

## Differences in Learner Behaviors Across Temperament Patterns

Data from sets of measures of the Student Behavior Observation Schedule that were collected at the three data collection points were combined and analyzed separately for the three temperament categories. The results are summarized in Table 3. Significant between-group differences were noted in students' interactions with teachers, in the initiation of activities, and in the manner in which time was spent. Students in the Easy and Difficult temperament categories initiated greater percentages of interactions with teachers than did their Slow-To-Warm-Up peers (70.5% and 62.5%, respectively, vs. 44%), and the Slow-To-Warm-Up students initiated their own learning activities least often (86.7% vs. 97.5% for the Easy students and 92.2% for the Difficult students). The Easy and Slow-To-Warm-Up students spent significantly greater amounts of time on-task (72.9% and 75.8%, respectively, vs. 66.1% for the Difficult students) and less time on distracted behavior (Easy, 24.3%; Slow-To-Warm-Up, 20.7%; Difficult, 30.2%). It is noteworthy that the general patterns of differences in student behaviors among the three groups were found for each of the three separate data collection periods.

## Differences in Learner Achievement Across Temperament Patterns

The differences in learner achievement across temperament patterns were examined based on task completion rates, mastery of curriculum objectives, and gains in scores on standardized achievement tests. Results of the comparisons are summarized in Table 4.

Task completion rates were defined as the ratios of numbers of tasks completed to numbers of tasks assigned. Mean task completion rates for students in each temperament category at each of the three data collection points, as well as the mean task completion rates for the entire school year, are indicated in Table 4. Overall, the data show that the Easy students had

Table 3

Mean Percentages of Observed Frequencies
of Learner Behaviors for Students in the
Three Temperament Categories

| | Temperament Category | | | |
|---|---|---|---|---|
| Learner Behavior Variable | Easy (N=23) | Difficult (N=13) | Slow-To-Warm-Up (N=20) | $F$- test of significant difference |
| Interactions | | | | |
| With Teachers | | | | |
| Initiation | | | | |
| Student | 70.5[a]<br>(4.42) | 62.5<br>(4.09) | 44.0<br>(4.48) | 3.25* |
| Teacher | 29.5<br>(4.42) | 37.5<br>(3.75) | 56.0<br>(5.60) | 3.25* |
| Purpose | | | | |
| Instructional | 83.3<br>(3.45) | 77.1<br>(3.33) | 77.6<br>(3.63) | |
| Management | 16.7<br>(3.45) | 22.9<br>(3.33) | 22.4<br>(3.63) | |
| With Other Students | | | | |
| Share Ideas | 98.0<br>(.84) | 94.8<br>(1.73) | 98.8<br>(.66) | |
| Disrupt | 2.0<br>(.84) | 5.2<br>(.52) | 1.2<br>(.66) | |
| Activity Type | | | | |
| Prescriptive | 77.2<br>(3.05) | 72.2<br>(3.36) | 70.7<br>(3.06) | |
| Exploratory | 22.8<br>(3.07) | 27.8<br>(3.24) | 29.3<br>(3.01) | |
| Setting | | | | |
| Group: interactive | 10.6<br>(2.22) | 8.7<br>(1.77) | 11.6<br>(2.17) | |
| Group: parallel | 35.2<br>(3.70) | 28.0<br>(3.24) | 37.5<br>(3.35) | |
| Individual | 54.2<br>(3.66) | 63.3<br>(3.52) | 50.9<br>(5.07) | |
| Initiation | | | | |
| Assigned | 2.5<br>(1.31) | 7.8<br>(1.62) | 13.3<br>(2.25) | 6.03* |
| Self-initiated | 97.5<br>(.55) | 92.2<br>(1.72) | 86.7<br>(2.25) | 8.57* |
| Manner | | | | |
| On-task | 72.9<br>(2.28) | 66.1<br>(2.34) | 75.8<br>(2.06) | 2.50* |
| Waiting for teacher help | 2.8<br>(.75) | 3.7<br>(.71) | 3.5<br>(.79) | |
| Distracted | 24.3<br>(2.35) | 30.2<br>(2.47) | 20.7<br>(2.05) | 2.21* |

Note. [a] Numbers in parentheses under mean percentages represent standard deviations.

* $p < .05$

Table 4

Summary of Learner Achievement Outcomes
for Students in the Three Temperament Categories

| | Temperament Category | | | |
|---|---|---|---|---|
| Learner Achievement Measure | Easy (N=23) | Difficult (N=13) | Slow-To-Warm-Up (N=20) | *F*-test of significant difference |
| Task Completion Rates[a] | | | | |
| September | 93.5% | 91.4% | 90.2% | .68 |
| December | 93.0% | 91.0% | 84.1% | 2.16* |
| May | 99.6% | 96.3% | 89.5% | 2.21* |
| Entire Year | 95.4% | 92.6% | 87.9% | 2.10* |
| Mean Number of Curriculum Objectives Mastered | | | | |
| Math | 27.76<br>(10.84)[b] | 30.83<br>(11.82) | 33.59<br>(10.80) | |
| Reading | 27.44<br>(11.12) | 26.67<br>(14.33) | 24.82<br>(11.72) | |
| Mean Percentile Rank on Metropolitan Achievement Test | | | | |
| Math | | | | |
| Fall | 75.217<br>(23.48) | 49.00<br>(37.07) | 70.62<br>(27.73) | |
| Spring | 94.850<br>(78.50) | 77.50<br>(24.31) | 82.38<br>(82.38) | |
| Gain | 19.63 | 28.50 | 11.76 | |
| Reading | | | | |
| Fall | 75.913<br>(20.45) | 49.833<br>(43.72) | 75.50<br>(32.05) | |
| Spring | 95.15<br>(11.18) | 78.25<br>(22.62) | 84.06<br>(24.14) | |
| Gain | 19.24 | 28.42 | 8.56 | |

Note.[a] Task completion rates are represented as the average percentage of assigned tasks that were completed correctly by students in each temperament category.

[b] Numbers in parentheses indicate standard deviations.

*$p < .05$

significantly greater task completion rates than did students in the other temperament groups. Task completion rates for the Easy students increased steadily throughout the school year. By contrast, an erratic pattern of task completion rates was found for the Slow-To-Warm-Up students.

No significant patterns of differences were found among the three groups in terms of the mean number of curriculum objectives mastered in reading and math. Although students in the Slow-To-Warm-Up category seemed to master the greatest mean number of math objectives and the least mean number of reading objectives, it is difficult to assess the meaning of these numbers. The number of objectives, and the level of coverage of the objectives, differ for various levels in the math and reading curricula.

To examine the differences in achievement across the three temperament categories, the mean percentile scores and the gain scores on the Metropolitan Achievement Test in reading and math were compared. As shown in Table 4, the results from an analysis of variance revealed several patterns of between-group differences that were consistent for math and reading. The fall achievement scores for the Easy and Slow-To-Warm-Up students were greater than those for the Difficult students. The spring scores indicate that the Easy students maintained reading and math scores higher than those of Difficult and Slow-To-Warm-Up students, and that the percentile rank scores for the Difficult and Slow-To-Warm-Up students were nearly comparable.

The similar patterns of differences in the gain scores among the three temperament categories also should be noted. The highest gain scores in reading and math were reported for the Difficult students, while the Slow-To-Warm-Up students had the lowest gain scores in both subject-matter areas.

## Differences in Learner Adjustment and Competence Across Temperament Patterns

Data from the Self-Responsibility Interview Schedule (SRIS), the Perceived Competence Scale (PCS) for children, and the Teacher Perception Questionnaire (TPQ) were analyzed to compare patterns of differences across the three temperament categories in terms of students′ and teachers′ perceptions of students′ adjustment and competence.

Results from analyses of the data obtained from the SRIS are summarized in Table 5. Three between-group differences were evident in the fall and spring data sets and in the total SRIS scores. On the questions which tap students′ knowledge of the learning environment, Easy students scored significantly higher than did students in the other temperament categories. Slow-To-Warm-Up students, on the other hand, scored the lowest of the three groups on this subscale of the SRIS. Regarding students′ perceptions of locus of control over their learning environment, the Easy students seemed to attribute the most personal control over their learning situation, while the scores for the Difficult students were strikingly low by contrast, indicating that these students attributed most of the control over their learning to external factors. No significant differences were found in students′ perceptions of their ability to evaluate their own learning. In the overall score, Easy students appeared to assume the greatest amount of responsibility

Table 5

Summary of Results From the Self-Responsibility Interview Schedule (SRIS)
for Students in the Three Temperament Categories

| | Subscale of the SRIS | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Knowledge of Learning Environment | | | | Locus of Control | | | | Evaluation of Work | | | | Total Score | | | |
| Temperament Category | Fall | | Spring | | Fall | | Spring | | Fall | | Spring | | Fall | | Spring | |
| | $\bar{X}$ | s.d. | $\bar{X}$ | s.d. | $\bar{X}$ | s.d. | $\bar{X}$ | s.d. | $\bar{X}$ | s.d. | $\bar{X}$ | s.d. | $\bar{X}$ | s.d. | $\bar{X}$ | s.d. |
| Easy (N = 23) | 4.33 | (.68) | 6.71 | (.49) | 3.65 | (.70) | 5.35 | (.51) | 5.56 | (.59) | 6.85 | (.63) | 12.72 | (.66) | 16.75 | (.49) |
| Difficult (N = 13) | 3.69 | (.89) | 5.76 | (.86) | 1.07 | (.99) | 1.25 | (.90) | 5.26 | (.45) | 6.90 | (.17) | 10.68 | (1.13) | 13.11 | (1.19) |
| Slow-To-Warm-Up (N=20) | 3.06 | (.89) | 5.57 | (.54) | 2.68 | (1.12) | 3.66 | (.87) | 5.08 | (.50) | 6.47 | (.55) | 11.75 | (.37) | 16.28 | (.28) |
| *F* - test of significant difference | 2.93* | | 2.45* | | 1.98* | | 3.63* | | 1.17 | | 1.88 | | 2.33* | | 3.47* | |

Note. $^{*}p \leq .05$

for managing their own learning and behavior, followed closely by the Slow-To-Warm-Up students. The Difficult students had significantly lower overall scores than either of the other groups.

Results from the fall and spring administrations of the PCS are summarized in Table 6. As shown in the table, the Slow-To-Warm-Up students tended to rate themselves lower in social competence in the fall than did either the Easy or Difficult students. However, by the spring administration of the PCS, their self-ratings of social competence were comparable with those of the other two groups. In fact, students in all three temperament categories showed increases in their self-ratings on the four subscales by spring. (The one exception was the Difficult students whose ratings on the subscale, social competence, were comparable in fall and spring.)

The analysis of teachers' responses on the TPQ revealed several trends. In general, teachers' ratings of the students tended to be higher in the spring than in the fall. The results reported in Table 7 show that for all the subscales of the TPQ, teachers rated most of the Easy students as "high." By contrast, students in the Difficult temperament category consistently were rated "low" by the teachers, with the exception of two subscales, independence in school learning and productivity or the amount of work they were capable of completing. On these subscales, the Slow-To-Warm-Up students received the greatest percentages of "low" ratings.

## Discussion

Overall, findings from the present study support the hypothesized relationship (see Figure 1) among temperament characteristics, specific

Table 6

Summary of Mean Rating Scores and Standard Deviations on the Perceived Competence Scale (PCS) for Children for Students in the Three Temperament Categories

| Temperament Category | Subscale of the PCS | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Cognitive Competence | | Social Competence | | Physical Competence | | Self-Esteem | |
| | Fall | Spring | Fall | Spring | Fall | Spring | Fall | Spring |
| Easy<br>(N=23) | 3.16<br>(.53)[a] | 3.53<br>(.53) | 3.40<br>(.59) | 3.57<br>(.39) | 2.95<br>(.78) | 3.09<br>(.82) | 3.31<br>(.48) | 3.65<br>(.34) |
| Difficult<br>(N=13) | 3.06<br>(.58) | 3.54<br>(.33) | 3.41<br>(.46) | 3.34<br>(.59) | 2.96<br>(.81) | 3.04<br>(.69) | 3.31<br>(.78) | 3.40<br>(.53) |
| Slow-To-Warm-Up<br>(N=20) | 2.88<br>(.63) | 3.45<br>(.41) | 2.70<br>(.80) | 3.04<br>(.71) | 2.84<br>(.39) | 2.96<br>(.81) | 3.07<br>(.61) | 3.41<br>(.61) |
| *F*- test of significant difference | 1.33 | .611 | 4.07* | 1.81 | .84 | .79 | .88 | 1.14 |

Note. [a] Numbers in parentheses indicate standard deviations.

*$p \leq .05$

**Percentages of Students in Each Temperament Category Rated as High, Average, and Low on the Subscales of the Teacher Perception Questionnaire (TPQ)**

| Subscale of the TPQ | Teachers' Rating of Students | Temperament Category | | | | | |
|---|---|---|---|---|---|---|---|
| | | Easy (N=23) | | Difficult (N=13) | | Slow-To-Warm-Up (N=20) | |
| | | N | % | N | % | N | % |
| Perceived Achievement<br>Which students do you believe are within the top 1/3, middle 1/3, and bottom 1/3 of your class, in terms of academic achievement? | High | (14) | 61 | ( 1) | 8 | ( 4) | 20 |
| | Average | ( 7) | 30 | ( 3) | 23 | ( 9) | 45 |
| | Low | ( 2) | 9 | ( 9) | 69 | ( 7) | 35 |
| Perceived Effort<br>Consider the effort students in your class display when working on an instructional task. Divide the students into three groups according to whether you believe they exert a lot of effort, an average amount of effort, or little effort in doing what they are expected to do in school. | High | (16) | 70 | ( 1) | 8 | ( 2) | 10 |
| | Average | ( 5) | 22 | ( 3) | 23 | (11) | 55 |
| | Low | ( 2) | 8 | ( 9) | 69 | ( 7) | 35 |
| Perceived Independence in School Learning<br>Which students do you believe are high, average, or low in their ability to successfully complete an instructional task independently, without teacher assistance? | High | (16) | 70 | ( 0) | 0 | ( 3) | 15 |
| | Average | ( 3) | 13 | ( 8) | 62 | ( 8) | 40 |
| | Low | ( 4) | 17 | ( 5) | 38 | ( 9) | 45 |
| Perceived Amount of Work Students Are Capable of Completing<br>Dividing students in your class into three groups, who do you believe is high, average, or low with respect to the number of instructional tasks they can complete? | High | (14) | 61 | ( 1) | 8 | ( 3) | 15 |
| | Average | ( 9) | 39 | ( 6) | 46 | ( 4) | 20 |
| | Low | ( 0) | 0 | ( 6) | 46 | (13) | 65 |
| Perceived Popularity of Students Within the Classroom<br>Which students are sought after by their classmates for instructional assistance most often, moderately often, and least often? | High | (16) | 70 | ( 0) | 0 | ( 3) | 15 |
| | Average | ( 6) | 26 | ( 4) | 31 | ( 9) | 45 |
| | Low | ( 1) | 4 | ( 9) | 69 | ( 8) | 40 |
| Perceived Control Over Student Learning Behavior<br>Think of the amount of influence you have over the learning outcomes of students in your class. Divide the students into three groups according to whether you believe you have a high, average, or small amount of influence over their learning outcomes. | High | (13) | 57 | ( 0) | 0 | ( 6) | 30 |
| | Average | ( 8) | 35 | ( 3) | 23 | ( 8) | 40 |
| | Low | ( 2) | 8 | (10) | 77 | ( 6) | 30 |

learning conditions, and learner behaviors and outcomes. Several distinct patterns of learner behaviors and learning outcomes for students with different temperament characteristics emerged from the study. Moreover, there is evidence to suggest that certain student characteristics may either facilitate or hinder certain student learning processes and outcomes under different environmental conditions (e.g., program design, teacher behaviors, teachers' and students' perceptions of the temperament characteristics of individual students). From a methodological perspective, findings from the study also seem to support the feasibility and meaningfulness of using triangulation procedures, observational data, student and teacher ratings, and achievement data in studying the interaction of these clusters of variables.

The importance of examining the relationship between student characteristics and classroom processes and outcomes in the context of the idiosyncracies of particular learning environments seems to be indicated by data from the present study. For example, the study was conducted in a classroom where there was strong emphasis on developing students' competence for taking increasing responsibility for their learning and behaviors. The data suggest that, in this setting, classroom processes such as student-initiated interactions with teachers and peers and student choice of learning activities emerged as more salient contributing variables in student achievement and adjustment than would be the case in traditional programs that are predominated by teacher-directed, group instruction. Thus, in a way, the data on the difference patterns in classroom processes for students with different temperament characteristics illustrate the interaction between student temperament and learning conditions in learning environments with different requirements.

The results for students in the Slow-To-Warm-Up category are another example of the role of the learning environment in affecting learner behaviors and outcomes. Of the three temperament categories, students in the Slow-To-Warm-Up group were least able to explain procedures, and they also rated themselves below their peers in social competence. Slow-To-Warm-Up students were observed to initiate a significantly lower percentage of interactions with the teachers, compared to the Easy and Difficult students, and to less frequently begin to work on learning tasks independently. In addition, teachers rated the Slow-To-Warm-Up students as low in independence and in the amount of instructional tasks that they could complete. Furthermore, the Slow-To-Warm-Up students made the smallest achievement gains. While the data do not provide information on the reasons for these findings, the ALEM′s emphasis on self-responsibility and student initiative, as well as the teachers′ practice of making instructional adaptations in accordance with perceived differences in learner characteristics in order to maximize each student′s chances to succeed in school learning, may have accentuated the problem of adjusting to classroom functioning for students in the Slow-To-Warm-Up temperament category. The Slow-To-Warm-Up temperament characteristics may have contributed to the difficulties these students experienced in adjusting to environments where student initiative is a functional requirement.

Furthermore, it can even be speculated that this problem was magnified by the ALEM's adaptive instruction approach. In fact, the data from the study that suggest teachers tended to assign fewer tasks to the Slow-To-Warm-Up students than were assigned to students in the Easy category, and to initiate a greater percentage of interactions with students in the Slow-To-Warm-Up category, may be a case in point. They illustrate the complexity and the interactive effects among the clusters of variables included in the conceptual model of temperament characteristics and school learning.

The findings for students in the Difficult temperament category also serve to highlight the importance of the interaction of learner characteristics with the learning environment. Difficult students were observed to spend less time on-task and to be more distracted than their classmates. This group also was rated by teachers as being highly distractible and low in persistence. They were rated most negatively by teachers on the learning outcomes, adjustment and competence, and they also were perceived as being less productive than the other two groups. Teachers indicated their feelings of less control over this group's learning behaviors than over those of the Easy and Slow-To-Warm-Up students. By contrast, it is interesting to note that the Difficult students' responses to the student perception measures suggested they felt the least in control of their own learning behaviors, compared to students in the other two groups. The data also suggest that the Difficult students were assigned fewer tasks and scored the lowest percentile ranks in reading and math on standarized achievement tests.

The instructional relevance of differences in students' temperament characteristics clearly is suggested by the findings from the present study and, consequently, implications can be drawn for improving classroom practice. The adjustments (intentional and unintentional) made by teachers in the amount of assignments for students in the three groups reflect their perceptions of differences in the students' temperament characteristics. The teachers also attempted to adapt their own behavior to the needs of students with different temperament dispositions. For example, teachers increased the initiation of learning tasks and assumed greater control over the learning behaviors of Slow-To-Warm-Up students who were reluctant to initiate interactions and learning activities. It can be hypothesized that students in the Slow-To-Warm-Up category were able to maintain high percentile ranks in their achievement scores and to develop increased perceptions of their own social competence in response to this teacher attention. Similarly, the positive student and teacher perceptions of the Easy students corresponded with the observations of these students to be highly task-oriented and persistent, and less distractible, in their learning behaviors. They were assigned more tasks and made the greatest achievement gains, compared to students in the other two groups.

To summarize, the study was initiated to increase understanding of the relationship among students' temperament characteristics and their learning conditions, learner behaviors, and learning outcomes. The data seem to suggest the dynamic and dialectical interaction among temperament characteristics and these clusters of variables. Moreover, the multiple and multidimensional interactive effects of certain temperament characteristics are evident. Temperament characteristics that tend to facilitate adaptation in some learning situations actually may be dysfunctional in others. In the present study, for example, the Slow-To-Warm-Up students were found to elicit more teacher-initiated interactions, and teachers tended to expect less work from, and assign less work to, Slow-To-Warm-Up and Difficult students.

Another instructional implication may be drawn from the distribution of greater percentages of Easy students in the higher grade levels. These results suggest the possibility of influencing temperament through consistent situational exposure. Just as learning environments can be adapted to individual learners, learner characteristics may be adaptable to the learning environment. The study points to the need for research to examine this continuous reciprocal adaptation in order to broaden understanding of the relationship among learner characteristics, learning conditions, learner behaviors, and learning outcomes. The possibility of improving schools' capacity to maximize each student's chances for success is determined, to a great extent, by the schools' capability to incorporate into their program design and implementation efforts effective responses to diverse student characteristics as well as modification of learners' abilities to meet the requirements of given learning tasks and to cope with the intellectual and social demands of the learning environment.

## Acknowledgment

The research reported herein was supported by the Learning Research and Development Center, which is supported in part as a research and development center by funds from the National Institute of Education (NIE) of the U. S. Department of Education. The opinions expressed do not necessarily reflect the positions or policies of NIE, and no official endorsement should be inferred.

# References

Berliner, D. C., & Rosenshine, B. V. (1976). The acquisition of knowledge in the classroom. Beginning Teacher Evaluation Study (Tech. Rep. IV-1). San Francisco: Far West Laboratory for Educational Research and Development.

Billman, J., & McDevitt, S. C. (1980). Convergence of parent and observer ratings of temperament with observations of peer interaction in nursery school. Child Development, 51, 385-400.

Bloom, B.S. (1968). Learning for mastery. Evaluation Comment, 1(2), 74-86.

Bloom, B.S. (1976). Human characteristics and school learning. New York: McGraw-Hill.

Brophy, J. E., & Good, T. L. (1974). Teacher-student relationships: Causes and consequences. New York: Holt, Rinehart and Winston.

Brophy, J., & Evertson, M. (1981). Student characteristics and teaching. New York: Longman.

Buss, A. H., & Plomin, R. A. (1975). Temperament theory of personality development. New York: Wiley.

Carey, W. B. (1981). The importance of temperament-environment interaction for child level and development. In M. Lewis & R. L. Rosenblum (Eds.), The uncommon child. New York: Plenum.

Carey, W. B., Fox, M., & McDevitt, S. (1977). Temperament as a factor in early school adjustments. Pediatrics, 60, 521-624.

Denham, C., & Lieberman, A. (Eds.). (1980). Time to learn. Washington, DC: National Institute of Education.

Dweck, C. S., & Goetz, F. E. (1977). Attributions and learned helplessness. In J. H. Harver, W. Ickes, & R. F. Kidd (Eds.), New directions in attribution research (Vol. 2). Hillsdale, NJ: Erlbaum.

Glaser, R. (1977). *Adaptive education: Individual diversity and learning.* New York: Holt, Rinehart and Winston.

Good, T. L., & Brophy, J. E. (1973). *Looking in classrooms.* New York: Harper and Row.

Gordon, E. W., & Wang, M. C. (1982). *Learner behavior and educational treatment: A multi-perspective analysis.* Learning Research and Development Center, University of Pittsburgh.

Graham, P., Rutter, M., & George, S. (1973). Temperamental characteristics as predictions of behavior disorders in children. *American Journal of Orthopsychiatry*, *43*(3), 328-339.

Hall, R. J., & Keogh, B. K. (1978). Qualitative characteristics of educationally high risk children. *Learning Disabilities Quarterly*, *1*(2), 62-68.

Harter, S. (1982). The Perceived Competence Scale for children. *Child Development*, *53*, 87-97.

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin*, *89*, 47-62.

Kagan, J., & Kogan, N. (1970). Individual variation in cognitive processes. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology* (Vol. I). New York: Wiley.

Keogh, B. K. (1982). Children's temperament and teachers' decisions. In R. Porter & G. M. Collins (Eds.), *Temperament differences in infants and young children* (Ciba Foundation Symposium No. 89). London: Pittman.

Keogh, B. K., Pullis, M. E., & Cadwell, J. (1982). *A short form of the Teacher Temperament Questionnaire*. Unpublished manuscript, University of California at Los Angeles.

Lambert, N. M., & Windmiller, M. (1977). An exploratory study of temperament traits in populations of children at risk. Journal of Special Education, 611(1), 37-47.

Marshall, H. H. (1981). Open classrooms: Has the term outlived its usefulness? Review of Educational Research, 51, 181-192.

Maurer, R., Cadoret, R. J., & Cain, C. (1980). Cluster analysis of childhood temperament data on adoptees. American Journal of Orthopsychiatry, 51, 522-534.

Meyer, W. T., & Thompson, G. G. (1966). Sex differences in the distribution of teacher approval and disapproval. Journal of Educational Psychology, 41, 385-396.

Mischel, W., Zeiss, R., & Zeiss, A. (1974). Internal-external control and persistence: Validation and implications of the Stanford-Preschool Internal-External Scale. Journal of Personality and Social Psychology, 29, 265-278.

Peterson, P. L. (1979). Direct instruction reconsidered. In P. L. Peterson & H. J. Walberg (Eds.), Research on teaching: Concepts, findings, and implications. Berkeley, CA: McCutchan.

Pullis, M. E. (1979). An investigation of the relationship between childrens' temperament and school adjustment. Unpublished doctoral dissertation, University of California at Los Angeles.

Rosenshine, B. V. (1979). Content, time, and direct instruction. In P. L. Peterson & H. J. Walberg (Eds.), Research on teaching: Concepts, findings, and implications. Berkeley, CA: McCutchan.

Rutter, M., Birch, H. J., Thomas, A., & Chess, S. (1964). Temperament characteristics in infancy and later development of behavioral disorders. British Journal of Psychiatry, 110, 651-661.

Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). Fifteen thousand hours: Secondary schools and their effects on children. Cambridge, MA: Harvard University Press.

Strom, C. D., & Wang, M. C. (1982, May). A validation study of a degree of program implementation assessment instrument. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.

Terestman, N. (1980). Mood quality and intensity in nursery school children as predictors of behavior disorders. American Journal of Orthopsychiatry, 50, 125-137.

Thomas, A., & Chess, S. (1977). Temperament and development. New York: Brunner/Mazel.

Thomas, A., Chess, S., & Birch, H. G. (1968). Temperament and behavior disorders in children. New York: New York University Press.

Thomas, A., Chess, S., Birch, H., Hertzig, M., & Korn, S. (1963). Behavioral individuality in early childhood. New York: New York University Press.

Wang, M. C. (1974a). The Self-Responsibility Interview Schedule. University of Pittsburgh, Learning Research and Development Center.

Wang, M. C. (Ed.). (1974b). The use of direct observation to study instructional-learning behaviors in school settings (LRDC Publication Series 1974/9). Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Wang, M. C. (1976). The use of observational data for formative evaluation of an instructional model. Instructional Science, 5, 365-369.

Wang, M. C. (1980a). Adaptive instruction: Building on diversity. Theory Into Practice, 19(2), 122-127.

Wang, M. C. (1980b). The degree of implementation assessment measures for the Adaptive Learning Environments Model. Experimental edition. University of Pittsburgh, Learning Research and Development Center.

Wang, M. C., Catalano, R., & Gromoll, E. (1983). Training manual for the Implementation Assessment Battery for Adaptive Instruction. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Wang, M. C., Gennari. P., & Waxman, H. C. (1985). The Adaptive Learning Environments Model: Design, implementation. and effects. In M. C. Wang & H. J. Walberg (Eds.), *Adapting instruction to student differences*. Berkeley, CA: McCutchan.

Wang, M. C., & Stiles, B. (1976). An investigation of children's concept of self-responsibility for their school learning. *American Educational Research Journal*, *13*(3), 159 179.

Weisstein, W.. & Wang, M. C. (1980, April). *An investigation of the effects of teacher perceptions on high and low achievers*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

# TEMPERAMENT CHARACTERISTICS AND LEARNING: COMMENTS AND CAVEATS

Samuel Messick

Educational Testing Service

To begin with, I would like to underscore two important messages of this symposium. First, all of the speakers take seriously the notion that personality characteristics may be important contributors to, or moderators of, learning and educational performance. In examining this possibility, the speakers also agree in highlighting selected dimensions of temperament or behavioral style -- variables which are more frequently perused in clinical settings in connection with behavior disorders than in educational settings in connection with learning processes and outcomes. Second, in attempting to relate temperament to learning, the speakers uniformly adopt a multivariate approach that embraces several dimensions of temperament operating conjointly, rather than focusing monolithically on a single broad dimension as has too often been the case in the study of cognitive styles (Messick, 1984).

However, on the down side, the dimensions of temperament included in these studies, being originally derived in psychiatric contexts (Thomas, Chess, & Birch, 1968), may not represent an optimal way of characterizing stylistic consistencies in behavior for educational purposes. These temperament dimensions may not be conceptualized or structured appropriately or at the right level of aggregation to reveal basic relationships to learning. Worse still, the commonly embraced multivariate approach may not be multivariate enough, both with respect to the range and variety of temperament consistencies relevant to education and with respect to nontemperament variables, such as intellective abilities, which should be controlled in appraising the impact of temperament on learning outcomes. I will return to these themes after first briefly addressing selected features of each presentation.

## Early Temperament and Later Educational Outcomes

In the longitudinal data reported by Lerner, Chess, and Lenerz, the only systematic outcome was the significant prediction of the reading, spelling, and arithmetic subscores of the Wide Range Achievement Test in adolescence by a composite of the five behavioral categories comprising the Easy versus Difficult temperament classification at age 5. The highest $R^2$ was .17 and the lowest was .10, which correspond to multiple R's of .41 and .32, respectively. This one systematic set of significant results out of multiple significance tests is only a modest finding, which becomes even less compelling when we realize that the reading, spelling, and arithmetic subscores are correlated and not independent criteria. However, it is important to note that these relationships were maintained with IQ controlled, so appropriate attention was paid to discounting a major rival interpretation. As an aside in connection with the prediction of learning outcomes from temperament, in the case vignettes appended to their study, we note that negative consequences of Difficult temperament are often overcome in the presence of high persistence, while positive expectations of Easy temperament are often contravened in the face of low persistence; so, perhaps the right combination of variables was not included in the regression analyses, or possibly persistence serves to moderate the temperament-performance predictions.

There were no significant relationships between early temperament and early adult educational attainment in the data reported. Furthermore, we note that the composite category of Difficult temperament is not very stable from childhood to adulthood (r = .15 from age 5 and .31 from age 3). This raises a question not only about the stability of the construct of Difficult temperament, but about its continuity as a construct over time. The structure of adult temperament may be different from that of child temperament, in which case a tracing of its developmental course becomes important -- or possibly the tracing of multiple developmental courses as a function of personality characteristics or of personality-environment interactions.

Since the structure of adult temperament may be more differentiated and complex than child temperament, the issue is to map changes in the number, nature, and interrelationships among dimensions and not just map changes in mean scores and variances. Such increased complexity of structure with age seems a distinct possibility when we review the structure of adolescent and adult temperament derived from self-reports and peer ratings as summarized, for example, by Guilford (1959) and Cattell (1979). An alternative strategy would be to start with the multiple dimensions of adult temperament and seek their roots in the existent child studies -- which would require a careful judgment as to whether the child studies included a sufficient range of behavioral variation to permit such a diverse array of roots to appear. An understanding of the course of development of temperament structures is especially important if predictive relationships are to be explored between early temperament and later temperament or between early temperament and later educational performance.

## Cross-Cultural Generalizability of Temperament Constructs

In the paper by Laosa, the finding of a disparate factor structure comprising two dimensions of mood/manageability and rhythmicity/responsivity in a Chicano sample at 3-1/2 and 4 years of age raises important questions of interpretability and methodology. To begin with, Laosa's sample of 100 children is quite small relative to the almost equal number (99) of behavioral variables, so caution must be exercised in the presence of likely structural instability. On the other hand, the analyses were carefully carried out in the manner of convergent triangulation, with important features being replicated from age 3-1/2 to age 4. Although Laosa is to be highly commended for using cross-age correlations among factor scores as one important basis for justifying the number of factors retained, it must be recognized that during a period of rapid developmental change, six months may be too long an interval to use as a criterion for short-term stability. By retaining only factors whose scores correlate substantially over that interval, one might miss genuine factors that are dropping out or coalescing with other factors during the first time period or are emerging or differentiating at the second time period. Of course, one would expect such factors to exhibit interpretable patterns of loadings at their respective times, and interpretability was another criterion emphasized by Laosa.

In addition to Laosa's general point that this two-factor structure looks different from those structures obtained in the Thomas-Chess tradition, there is the finer point that the nine Thomas-Chess behavioral categories break down in that in each instance some of the behaviors in each category relate to one of Laosa's factors and some to the other factor. This raises the question of the robustness of the Thomas-Chess behavioral categories at the level of basic temperament dimensions -- which, of course, also has ramifications for higher-order classifications that are derived from the basic nine, such as Easy versus Difficult temperament. Furthermore, it is not so much the presence of Laosa's data that calls that robustness into question as it is the absence of compelling analyses of dimensional structure -- from behaviors to behavioral categories or dimensions to higher-order classifications -- at different focused age levels and in different population groups.

The point about focused age levels is a potentially important one. Although the two factors obtained by Laosa correlated highly between age 3-1/2 and age 4 and the patterns of factor loadings were similar in tapping overlapping behavioral domains, they were by no means identical. The question is whether the obtained differences reflect random fluctuation or real development. If the latter, then studies pooling children who vary in age by 3 or 4 years or more may yield seriously distorted conglomerate structures that do not represent anyone very well.

With respect to interpreting the sources or determinants of Laosa's disparate factor structure, the situation is murky indeed in the absence of a comparison group or groups to control for extraneous factors. Without such controls, we cannot tell whether the disparate structure is attributable to being Chicano, attributable to the specific sociodemographic status of the population under study apart from being Chicano, attributable to the age range of the subjects studied, or to the unique oral-administration testing conditions employed. But one thing is certain, the issue of population validity -- of the generalizability of temperament structures across different population groups, of the meaning of temperament measures across population groups -- is clearly raised by Laosa's arguments if less clearly by his data. The data are equivocal in some respects, to be sure, but the issues may be ignored only at our peril.

## Influence of Children's Temperament on Their Personal-Social and Educational Experiences in School

In the studies reviewed by Pullis, the main feature I would like to highlight is that temperament is measured directly, not in terms of the nine Thomas-Chess behavioral categories, but in terms of the three factors of the Teacher Temperament Questionnaire (TTQ) derived from them. These factors comprise task orientation, which subsumes the Thomas-Chess categories of activity, persistence, and distractibility; personal-social flexibility, which subsumes adaptability, approach-withdrawal, and positive mood; and reactivity, which subsumes threshold, response intensity, and negative mood. Rhythmicity does not appear in this factor structure. The issue of population validity again arises because tacit in the report of the experimental findings is the presumption that this structure holds for moderately handicapped, emotionally disturbed, and learning disabled students.

# Differences in Student Temperament Characteristics and Their Effects on Classroom Processes and Outcomes

In the paper by DeStefano, Wang, and Gordon, the authors are to be simultaneously commended and chided on a number of counts. On the one hand, they do not merely presume that the three-factor structure of the TTQ holds for their primary school sample -- they perform a new factor analysis. On the other hand, they include in their factor analysis not only the individual behavioral items but the three subscale scores, the correlated errors from which are just about guaranteed to yield the expected three factors as statistical artifacts. On the one hand, they insist that researchers must address the high probability of changes in patterns of temperament characteristics as a function of experience and development. On the other hand, they pool students from 4 to 9 years of age to yield a single conglomerate temperament structure. On the one hand, they ask what identifiable patterns of temperament characteristics are present within this sample of elementary school-age children. On the other hand, they restrict attention to only the three modal profiles derived from the Thomas-Chess program -- namely, Easy temperament, Difficult temperament, and Slow-To-Warm temperament. These modal profiles are defined across the three TTQ factors of task orientation, personal-social flexibility, and reactivity.

Several questions arise that cannot be addressed in depth here. For example, if the three-factor structure can be shown to hold in this sample in the absence of correlated errors (i.e., by dropping the three subscale scores from the factor analysis), are there other modal profiles that can be empirically derived with these data? Furthermore, once temperament patterns are to be related to learner achievement and competence outcomes, we must inquire about the need for controls for intellective abilities and prior achievements. This is especially important since some purported temperament constructs such as task orientation and adaptability may overlap both conceptually and behaviorally with some cognitive constructs of intellective ability. It will be recalled that Binet emphasized in his conceptualization of intelligence not only comprehension, reasoning, and judgment but also attention-control and directed thinking. Both in the interests of parsimony and on theoretical grounds, we should relate temperament not to learner performance directly, but to residual performance variance once intellective determinants are controlled for.

## Overview

The cumulative message in these papers -- apart from the need for controls for intellective abilities and prior achievements in relating temperament characteristics to learning -- is the need for careful attention to the measurement and structural analysis of temperament, especially with regard to the issue of population validity and the possibility of developmental changes in temperament structure. Given the possibility of different structures for different groups and of changes in structure with development, it might be wise to avoid prematurely tying measurement to derived structures such as the three factors of the TTQ or to specific modal profiles such as Easy and Difficult temperament. It might be better to measure temperament in terms of homogeneous and reliable behavioral dimensions such as the nine Thomas-Chess categories, if they hold up, and to use group differences in the correlations among these behavioral dimensions as signals of possible differences in the derived structures for the groups under investigation.

It would also be beneficial to investigate modal profiles defined across the basic behavioral dimensions rather than across derived factors combining these dimensions. There are two important reasons for recommending this strategy. One reason is that the broader array of behavioral dimensions relative to the smaller number of derived factors offers greater opportunity and richness for the emergence of empirically derived modal profiles. Second, when related to classroom processes and learning outcomes, these modal temperament profiles are closer to the data, easier to interpret, and less likely to be misconstrued and misapplied vis-à-vis needed controls.

# References

Cattell, R. B. (1979). *Personality and learning theory (Volume I): The structure of personality in its environment*. New York: Springer.

Guilford, J. P. (1959). *Personality*. New York: McGraw-Hill.

Messick, S. (1984). The nature of cognitive styles: Problems and promise in educational practice. *Educational Psychologist*, *19*, 59-74.

Thomas, A., Chess, S., & Birch, H. G. (1968). *Temperament and behavior disorders in children*. New York: New York University Press.

# TEMPERAMENT AND LEARNING: A COMMENTARY

Edmund W. Gordon

Yale University

It must be gratifying, and at the same time frustrating, to investigators who have broken new ground and have become identified with a widely held conceptual paradigm, to see their influence on a field within their lifetimes. It must be gratifying to Drs. Chess and Thomas to see so many of us recognizing the importance and utility of their nine categories of temperament in our continuing efforts at understanding human behavior. It must be superficially flattering but frustrating to note that the conceptual paradigm, in large measure, is accepted uncritically and is without serious challenge. Chess and Thomas have acknowledged this ambivalance in personal communication indicating that, while they have identified and explicated the conceptual model, it is up to younger workers to validate or discard, challenge and refine. Since they are in our presence at this symposium, we acknowledge our debt to them and their late colleague Herbert Birch, and we honor them with these replications as well as implicit challenges to their work.

The papers by DeStefano, Wang, and Gordon; Laosa; Lerner, Chess and Lenerz; and Pullis rest on the assumption that temperament is an important multidimensional component of personality. Three of these papers examine the relationship of temperament to directed learning (education). One examines the impact of cognitive competence and selected family background characteristics of temperament. Dr. Messick has already complimented the authors of these papers even as he criticized them on methodological and theoretical grounds. I too found the papers to be interesting and thoughtful extensions of the growing body of important work. Yet these papers provoke interesting questions, the resolutions of which may be essential preconditions for the further advancement of the Chess and Thomas paradigm. Among these are such issues as:

1. the nature and validity of the construct itself;

2. the stability and consistency of its expression;

3. the impact of these behavior tendencies on the adaptive performance of the learning person;

4. the impact of these behavior tendencies on other persons.

With respect to the nature and validity of the construct, there is a problem of what validity can be established for what are essentially subjective descriptive identifications of behavior. Thomas, Chess, and Birch (1968) define temperament as "a phenomenologic term used to describe the characteristic tempo, rhythmicity, adaptability, energy expenditure, mood and focus of attention" of an individual, factors thought to function "independently of the content of any specific behavior" (p. 4). In his work, Catell (1950) has asserted that temperament includes those traits that are uninfluenced by incentive or complexity, and he illustrates the phenomenon with such characteristics as high strungness, speed, energy, and emotional reactivity, which he argued, "common observation suggests are largely constitutional" (p. 35). Both of these definitions depend heavily on the iteration of terms descriptive of the phenomenon being defined. Such definition leaves them conceptually imprecise and subject to observer attribution. The Chess, Thomas, and Birch use of the term "characteristic" to mean typical, and the Catell assertion of a constitutional basis for the phenomenon, lay the foundation for the colloquial view that the behavior is stable and may be consistently manifested in the behavior of the individual. Thomas and Chess (1980) subsequently have challenged this colloquial view, but this commentator (Gordon, 1973) has argued in support of the possible constitutional basis of the behavior, based upon having observed with Birch early precursors of temperament in neonates. This confusion with respect to definition and nature of temperament may be related to the rather primitive stage of our technology for assessing temperament. We continue to rely on observations and judgement by persons who are assumed to know the subject well and to have observed over a variety of situations. However, in the procedures for the assessment of the behavior, even in the presence of high reliability among observers, we are at a loss to determine whether we are dealing with constancy of the expression of a behavior or constancy of the stimulus conditions which evoke a particular behavior. One could argue that, for a characteristic to be intrinsic to the person, it should be constant in its expression across different situations. We have not seen studies of temperament in which the primary features of situations have been deliberately varied and manifestations of temperament monitored. Even in such a test, the veridicality of the phenomenon remains at issue. When we add problems already identified by Messick (i.e., the overlap between categories of temperament and the confounding of some of the categories by intellect), we are left with a construct that has considerable conceptual appeal but whose nature and validity are unclear.

Thomas and Chess (1980) have identified the question of the stability and consistency of the construct as an important one. In the papers presented today, our authors implicitly assume that they are dealing with behaviors that are stable and consistent, at least over relatively short periods of time. There are two subproblems here. One has to do with a question of whether it is the behavior or the perception of the behavior which is consistent. We know that attribution phenomena confound all observations. The second subproblem is even more conceptually troublesome. Thomas and Chess (1980) make the point very clearly.

> As we ourselves originally began to observe clinically and impressionistically the phenomenon of temperament, we were struck by the many dramatic evidences of continuity in individuals we knew, sometimes from early childhood to adulthood. It was tempting to generalize from these instances to the concept that an older child or adult's temperamental characteristics could be predicted from a knowledge of his behavior style in early childhood. However, such a formulation would have been at complete variance with our fundamental commitment to an interactionist viewpoint, in which individual behavioral development is conceived as a constantly evolving and changing process of organism-environment interaction. (pp. 123-124)

The authors go on to present, in summary fashion, inconclusive clinical and empirical data but draw conclusions consistent with their belief. We are left by their work, as well as that of the authors at this symposium, with unresolved questions concerning the stability and consistency of the construct as manifested in behavior. It appears that where developmental demands and conditions are stable over time, the manifestation of temperament is relatively consistent. However, like most other behavioral phenomena, when the context is unstable and subject to significant changes, the characteristics shaped by such encounters are not consistent. It is also important to note that since the nine categories of temperament are not all of the same order, some may be more consistent or labile than others. Given these possibilities, the assumption of stability, which seems implicit in the four papers presented today, may be problematic.

There is a long history of trait-treatment-interaction studies in which questions related to the impact of a trait on the performance of the individual in a specific treatment have been studied. This work is largely equivocal (Cronbach & Snow, 1977), which is in part the case because the search is usually for interaction effects between single variables and single treatments. In the studies reported at this symposium, efforts have been made to use higher order independent variable clusters (i.e., combinations or orchestrations of traits, to produce meta-traits, viz., "Easy to Warm," "Difficult," "task orientation," "personal-social flexibility"). Despite some problems in the genesis of these meta-traits, the conceptualizations upon which they rest are on target. The adaptive behavior on humans is almost never univariant. Certainly, to understand complex behaviors, we must find ways to capture the dynamic patterning of the components of organized behavioral responses such as the impact of personal characteristics on human peformance.

Not only does the patterning of personal characteristics affect performance, it also affects the way in which the performer is responded to. Two of the papers presented build on the earlier work by Keogh, Pullis, and Cadwell (1982), which clearly indicates that one of the important aspects of temperament has to do with its role in influencing how teachers treat students, including the amount and kind of attention given and the way in which they evaluate student performance.

# References

Catell, R. B. (1950). *Personality: A systematic and factual study.* New York: McGraw-Hill.

Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods.* New York: Irvington.

Gordon, E. W. (1973). Affective response tendencies and self-understanding. *Proceedings of the 1973 Invitational Conference on Testing.* Princeton: Educational Testing Service.

Keogh, B., Pullis, M., & Cadwell, J. (1982). A short form of the Teacher Temperament Questionnaire. *Journal of Educational Measurement, 19,* 323-329.

Thomas, A., & Chess, S. (1980). *The dynamics of psychological development.* New York: Brunner/Mazel.

Thomas, A., Chess, S., & Birch, H. (1968). *Temperament and behavior disorders in children.* New York: New York University Press.